Evaluación de programas

Catalina Martínez Mediano

Evaluación de programas

CATALINA MARTÍNEZ MEDIANO

UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA

EVALUACIÓN DE PROGRAMAS

Quedan rigurosamente prohibidas, sin la autorización escrita de los titulares del Copyright, bajo las sanciones establecidas en las leyes, la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, y la distribución de ejemplares de ella mediante alquiler o préstamo públicos.

© *Universidad Nacional de Educación a Distancia*
Madrid 2017

www.uned.es/publicaciones

© *Catalina Martínez Mediano*

ISBN electrónico: 978-84-362-7307-6

Edición digital: noviembre de 2017

ÍNDICE

Presentación de la asignatura de evaluación de programas

1. Introducción
2 Objetivos de la asignatura contextualizados en la carrera del grado en Pedagogía
3. Requisitos para cursar la asignatura
4. Resultados de aprendizaje
5. Orientaciones al estudio

Tema 1. **La evaluación de programas educativos. Teoría de la evaluación de programas. Concepto y funciones de la evaluación de programas**

1. Objetivos y resultados de aprendizaje
2. Introducción
3. Introducción a la evaluación de programas
   3.1. La evaluación como disciplina de conocimiento
   3.2. La evaluación de programas y la investigación evaluativa
   3.3. Los métodos en la evaluación de programas
   3.4. De los métodos a la teoría de la evaluación de programas
   3.5. La naturaleza de la teoría de programas y de la evaluación de programas
4. El concepto de evaluación de programas y los criterios y las referencias para la evaluación
   4.1. Concepto de evaluación de programas
   4.2. Los criterios y las referencias para la evaluación de programas
5. Las funciones de la evaluación de programas
   5.1. La evaluación formativa y sumativa
   5.2. La evaluabilidad y su aplicación para la evaluación de los programas

5.3. La meta-evaluación y su función formativa y sumativa
5.4. La evaluación para el rendimiento de cuentas

6. El objeto de la evaluación
6.1. La evaluación de los aprendizajes de los estudiantes
6.2. La evaluación de la calidad de los centros educativos y conceptos específicos
6.2.1. La evaluación de la gestión de la calidad de los centros educativos. Los principios de la gestión de la calidad total Las organizaciones que aprenden
6.2.2. El aseguramiento de la calidad
6.2.3. Evaluación de programas y auditoría
6.2.4. Funciones del evaluador según su relación con la organización evaluada y el enfoque de evaluación utilizado

7. Propósitos de la evaluación de programas

Ejercicios de autoevaluación

Referencias bibliográficas

Tema 2. **Michael Scriven. La evaluación como la ciencia de la valoración. La evaluación orientada al consumidor. La lista de las claves de control para la evaluación y la meta-evaluación**

1. Objetivos y resultados de aprendizaje

2. Introducción

3. La teoría de la evaluación de Scriven
3.1. La evaluación como la ciencia del valor
3.2. Las necesidades sociales y la ciencia de la valoración
3.3. La evaluación libre de costes
3.4. La evitación de los prejuicios
3.5. La evaluación libre de metas
3.6. La evaluación formativa y sumativa
3.7. La meta-evaluación
3.8. La lógica de la evaluación y la práctica evaluativa

4. La evaluación orientada al consumidor

5. La lista de las claves de control para la evaluación y la meta-evaluación

6. Tabla resumen del enfoque de evaluación de programas de Scriven

Ejercicios de autoevaluación

Referencias bibliográficas

Tema 3. **El modelo CIPP de Daniel Stufflebeam: La evaluacion orientada a la toma de decisiones para la mejora de los programas y de las organizaciones**

1. Objetivos y resultados de aprendizaje

2. Introducción

3. Antecedentes del modelo CIPP. Las recomendaciones del Comité PDK

4. La teoría evaluativa y los tipos de decisiones que orienta la evaluación en el modelo CIPP
   4.1. El concepto de evaluación de programas del modelo CIPP
   4.2. Los cuatro tipos de evaluación del modelo CIPP
   4.3. Funciones formativa y sumativa de la evaluación de contexto, de entrada, de proceso y de producto
   4.4. Componentes esenciales de valor del modelo CIPP
   4.5. Filosofía y códigos éticos del modelo CIPP
   4.6. El modelo CIPP como un sistema para la mejora de los programas y de las organizaciones

5. La práctica evaluativa y las decisiones que orientan los tipos de evaluación del modelo CIPP
   5.1. La evaluación del contexto. Decisiones de planificación
   5.2. La evaluación de entrada. Decisiones de estructura
   5.3. La evaluación del proceso. Decisiones de realización
   5.4. La evaluación del producto. Decisiones de reciclaje

6. La lista de control para la evaluación

7. Tabla resumen del modelo CIPP de evaluación de programas de Stufflebeam

Ejercicios de autoevaluación

Referencias bibliográficas

Tema 4. **Ramón Pérez Juste: La evaluación de programas educativos**

1. Objetivos y resultados de aprendizaje
2. Introducción
3. La teoría evaluativa
   - 3.1. Educación y evaluación
   - 3.2. La evaluación integral, integrada e integradora
   - 3.3. El concepto de evaluación de programas
   - 3.4. Características de los programas educativos
   - 3.5. Función esencial de la evaluación: la mejora
4. El modelo evaluativo para la práctica evaluativa
   - 4.1. Etapa inicial: Evaluación del programa en sí mismo
     - 4.1.1. La calidad intrínseca del programa
     - 4.1.2. La adecuación al contexto
     - 4.1.3. Adecuación a la situación de partida
     - 4.1.4. Criterios de evaluación
   - 4.2. Etapa procesual: Evaluación del desarrollo de los programas
     - 4.2.1. Criterios para la evaluación de la implementación de los programas
     - 4.2.2. Metodología adecuada a la evaluación de procesos
   - 4.3. Etapa final: Evaluación de los resultados de los programas
     - 4.3.1. El programa en sus logros
     - 4.3.2. Críterios y referencias de la evaluación de resultados de la aplicación del programa
     - 4.3.3. La valoración de los resultados
   - 4.4. Institucionalización de la evaluación de programas
5. La pauta para la evaluación de programas educativos
6. Tabla resumen del modelo de evaluación de programas de Pérez Juste

Ejercicios de autoevaluación

Referencias bibliográficas

TEMA 5. **LOS MODELOS LÓGICOS COMO MARCO TEÓRICO Y DE ORGANIZACIÓN PARA EL DISEÑO, APLICACIÓN Y EVALUACIÓN DE PROGRAMAS EDUCATIVOS**

1. Objetivos y resultados de aprendizaje
2. Introducción
3. Los modelos lógicos para el diseño y evaluación de programas
   - 3.1. Los modelos lógicos como secuencia para atender a unas necesidades
   - 3.2. Marco teórico y de organización orientado a la consecución de los resultados
4. La evaluación del programa
5. Los criterios de calidad
6. Las metodologías de investigación para la evaluación
   - 6.1. Las metodologías de investigación mixtas para atender la complejidad
   - 6.2. La elaboración de instrumentos para la recogida de información
   - 6.3. Compromiso con el rigor y la responsabilidad en la investigación evaluativa
7. Elaboración de los informes de la evaluación de programas
8. Utilización de los resultados de la evaluación
9. Conclusiones

Ejercicios de autoevaluación

Referencias bibliográficas

TEMA 6. **LAS NORMAS DE CALIDAD DEL COMITÉ CONJUNTO PARA LA EVALUACIÓN DE PROGRAMAS EDUCATIVOS**

1. Objetivos y resultados de aprendizaje
2. Introducción
3. Características y conceptos de los estándares de calidad de la evaluación de programas
   - 3.1. Características de las Normas en su tercera edición

3.2. Los responsables de los programas en los centros educativos
3.3. Conceptos y definiciones sobre los estándares para la evaluación de programas

4. Las Normas para la evaluación de programas del Joint Committee
4.1. Normas de utilidad
4.2. Normas de viabilidad o factibilidad
4.3. Normas de probidad, o adecuación legal y ética
4.5. Normas de responsabilidad de la evaluación

5. Utilización de las Normas del Joint Committee para el desarrollo de los programas, de su evaluación y de su responsabilidad pública

6. Lista de control de las Normas para la comprobación de su aplicación en la evaluación de los programas

Ejercicios de autoevaluación

Referencias bibliográficas

Tema 7. **Técnicas e instrumentos para la recogida de información en la evaluación de programas**

1. Objetivos y resultados de aprendizaje

2. Introducción

3. La observación
3.1. La observación sistemática y la observación participante
3.2. Los sistemas de recogida de información en la observación
3.3. Cuestiones sobre el muestreo y las fuentes de error en la observación
3.4. Instrumentos para el registro de la observación
3.5. Características de la observación participante
3.5.1. La función del observador en la observación participante
3.5.2. El diseño para la realización de la observación participante
3.5.3. El proceso de recogida de datos en la observación participante
3.5.4. El análisis de la información en la observación participante
3.5.5. La elaboración del informe de la observación participante

4. La evaluación de necesidades

5. El cuestionario

6. La entrevista

7. El análisis de contenido
   7.1. Concepto y alcance del análisis de contenido
   7.2. Las fases del análisis de contenido
   7.3. Las unidades de análisis y su contexto
   7.4. La inferencia como objetivo intelectual del análisis de contenido y la validez como criterio de éxito

8. La técnica Delphi

9. Los grupos de discusión

Ejercicios de autoevaluación

Referencias bibliográficas

Tema 8. **Técnicas para el análisis de la información en la evaluación de programas**

1. Objetivos y resultados de aprendizaje

2. El análisis de los datos en la investigación cualitativa
   2.1. Descubrimiento en progreso
   2.2. Codificación de los datos
   2.3. Relativización de los datos

3. Criterios para valorar las investigaciones cualitativas
   3.1. Criterios de credibilidad, seguridad, confirmabilidad y transferibilidad
   3.2. Criterios hermenéuticos
   3.3. Criterios de autenticidad

4. El análisis de datos en la investigación cuantitativa
   4.1. Distribuciones numéricas y reducción de datos
   4.2. Niveles de medida y tipos de variables
   4.3. La distribución continua y la distribución normal de probabilidad
   4.4. Representaciones gráficas
   4.5. Medidas de tendencia central. La moda, la mediana y la media

4.6. Medidas de posición. Los cuartiles, deciles, percentiles
4.7. Medidas de variabilidad. La amplitud, la desviación media, la varianza y la desviación típica
4.8. Puntuaciones individuales. Las puntuaciones diferenciales y las puntuaciones típicas

5. Las medidas correlacionales bivariadas y su interpretación
5.1. La correlación producto momento de Pearson
5.2. La correlación ordinal o de rangos de Spearman

6. Comparación entre grupos. Los diseños de grupo
6.1. Concepto de experimento y de diseño de grupo experimental
6.2. Resolución de un diseño investigación de dos grupos independientes y nivel de medida de intervalo
6.3. Resolución de un diseño investigación de dos grupos relacionados y nivel de medida de intervalo

Ejercicios de autoevaluación

Referencias bibliográficas

ANEXO
ORIENTACIONES PARA EL ESTUDIO Y LA REALIZACIÓN DE UN TRABAJO PRÁCTICO

I. **Prueba de evaluación continua**

II. **Orientaciones para la realización de un trabajo práctico sobre la evaluación de un programa educativo**

1. Objetivos y resultados de aprendizaje

2. Introducción

3. Estructura del trabajo
3.1. Resumen
3.2. Introducción
3.3. El programa a evaluar
3.3.1. Fundamentos teóricos del programa
3.3.2. Objetivos del programa
3.3.3. Sujetos a los que se dirige el programa

3.3.4. El contexto
3.3.5. Recursos
3.4. La aplicación del programa
3.4.1. Estrategias y actividades
3.4.2. Utilización de los recursos
3.4.3. Resultados inmediatos a conseguir
3.4.4. Resultados a largo plazo
3.5. La evaluación del programa
3.5.1. Los objetivos de la evaluación, finalidad de la evaluación del programa
3.5.2. Funciones evaluativas
3.5.3. Modelo evaluativo elegido y Fases de la aplicación del modelo
3.5.4. Metodología de la evaluación, técnicas de recogida y análisis de datos
3.6. Resultados de la evaluación del programa
3.6.1. Presentación y valoración de los resultados
3.6.2. Conclusiones, discusión, valoración y propuestas de mejora
3.7. Aplicación de los estándares de la evaluación de programas
3.7.1. Aplicación de las normas a la evaluación
3.7.2. Valoración del cumplimiento de los estándares
3.8. Referencias bibliográficas y sitios webs
3.9. Anexos

4. Recomendaciones de estilo para la elaboración de trabajos de acuerdo con las normas APA

5. Orientaciones para la realización del trabajo práctico y criterios para su evaluación

Referencias bibliográficas

Aquí podrá encontrar información adicional
y actualizada de esta publicación

PRESENTACIÓN DE LA ASIGNATURA DE EVALUACIÓN DE PROGRAMAS

1. INTRODUCCIÓN

En las últimas tres décadas, los modelos de evaluación de programas con objetivos de mejora y de aseguramiento de la calidad para atender a las necesidades y problemas educativos y sociales, se han convertido en elementos esenciales para dirigir los esfuerzos por la mejora continua de los programas y de las organizaciones. Todo parece ser motivo de evaluación: el aprendizaje de los alumnos, el desempeño docente, el diseño y desarrollo del currículo y de los programas educativos, la organización de centros, entre otros. Los propósitos son muy diversos. Se puede evaluar para conocer, para valorar, para promocionar, para mejorar, pero también para rendir cuentas ante la sociedad, para acreditar o para decidir el futuro de un programa educativo. El salto cualitativo de la evaluación es su orientación hacia la mejora, en su función formativa, de modo que, entre todos los propósitos que pueden inspirar la tarea evaluativa, merece la pena destacar el de la mejora, ya sea referida a la formación del estudiante, del docente, de la organización, o de los sistemas educativos. La evaluación para la mejora se constituye en una dimensión irrenunciable de la tarea del evaluador.

Las intervenciones sobre las personas y las instituciones suponen una intención de mejora. Intervenimos en la sociedad a través del desarrollo de programas *ad hoc* para mejorar la realidad. Pero para mejorar no es suficiente con intervenir; tenemos que comprobar si nuestra intervención produce los cambios y mejoras deseados, pretendidos. Y para ello necesitamos de la evaluación.

La evaluación nos ayuda a recoger información sobre las necesidades que fundamentan nuestra intervención y sobre los programas y acciones que desarrollamos para atender a esas necesidades. Nos ayuda a recoger información para conocer cómo se aplican dichos programas y los cambios inmediatos y en el medio y largo plazo que producen en las personas, en las organizaciones y en la sociedad, con el objetivo de utilizar di-

cha información para introducir los cambios necesarios en los programas, los contextos, la formación, los procedimientos de intervención para conseguir las mejoras deseadas.

Se evalúa para mejorar la calidad de nuestros programas y de nuestras intervenciones, los contextos, la formación del personal implicado en los programas, de aquéllos que los aplican y de quienes los reciben, de las instituciones y de la sociedad que se beneficia directa o indirectamente de ello. La evaluación es una herramienta al servicio de la mejora de la calidad.

Los procedimientos que utiliza la evaluación están al servicio de los objetivos de las intervenciones y a ellos deben adaptarse. La realidad social es compleja, también sus necesidades y los programas que se desarrollan para atenderlas. También son complejos los contextos en los que se aplican, los resultados que producen y su impacto, incluidos los efectos por transferencia a otras situaciones y personas. Esta complejidad hace necesaria la adecuación de los procedimientos para la evaluación, la utilización de las metodologías de investigación de un modo riguroso, adaptadas a los múltiples objetivos que abarca la evaluación en los contextos en los que los programas y sus aplicaciones se desarrollan. Por todo ello la evaluación no es una disciplina sencilla, sino compleja, que requiere estudio, profundización, reflexión, interacción con la realidad, con los programas, en sus contextos, con las personas responsables de los mismos y con las que están interesadas en saber como funcionan y qué resultados consiguen para mejorar las intervenciones en un ciclo de mejora continua. La evaluación siempre aportará información para la mejora, para introducir cambios y actualizaciones, para adaptarse a las nuevas demandas de los avances que continuamente se producen en la sociedad.

2. OBJETIVOS DE LA ASIGNATURA CONTEXTUALIZADOS EN LA CARRERA DEL GRADO EN PEDAGOGÍA

La asignatura de evaluación de programas pretende capacitar a los alumnos de la carrera de Pedagogía en el desempeño de las siguientes competencias profesionales:

— Planificación de evaluaciones de programas.

— Realización de evaluaciones de programas.

— Elaboración de propuestas de cambio y mejora de los programas.

— Valoración crítica de informes de evaluación de programas.

— Búsqueda de nuevos conocimientos acerca de nuevos modelos y experiencias evaluativas.

Dado que es una asignatura que se imparte en el último curso del Grado en Pedagogía, los estudiantes habrán podido cursar la mayoría de las asignaturas del grado y contarán con un bagaje en conocimientos fundamentales sobre los contenidos del Grado de Pedagogía, tanto desde una perspectiva teórica como metodológica y aplicada, así como experiencia en la realización de trabajos para la comprensión de la relación entre teoría y práctica, que les permitirá aprovechar muy satisfactoriamente los contenidos de la asignatura y su aplicación para su desarrollo futuro como profesionales de la Educación. Todo ello contribuirá a una mejor comprensión de la asignatura de Evaluación de Programas.

3. REQUISITOS PARA CURSAR LA ASIGNATURA

El plan de estudios no establece ningún requisito formal previo para cursar las distintas asignaturas que lo integran.

Sin embargo, el aprendizaje en la asignatura de Evaluación de Programas se beneficia de los conocimientos de asignaturas tales como Métodos de investigación, Técnicas e instrumentos de recogida de información, Técnicas para el análisis de datos apoyadas en la Estadística y Evaluación del aprendizaje de los estudiantes. Otras asignaturas pertenecientes a otras materias también contribuyen a su aprovechamiento, como las asignaturas de Organización y gestión de centros educativos, la Didáctica, o la Elaboración de planes y programas, por citar sólo algunas.

De un modo más concreto, mencionamos los siguientes requisitos para cursar esta asignatura:

— Conocimientos para la utilización de ordenadores, conocimiento de procesadores de texto.

— Acceso a un ordenador con conexión a internet.

— Correcta expresión escrita, tanto en lo referente a la sintaxis como a la ortografía.

4. RESULTADOS DE APRENDIZAJE

Como resultado principal, se espera que los alumnos consigan las Competencias Específicas (CE) de la asignatura, que son:

— Desarrollo de competencias profesionales para la evaluación de programas.

— Realización de estudios prospectivos y evaluativos sobre características, necesidades y demandas pedagógicas, que originarán la información necesaria para que el pedagogo profesional desarrolle intervenciones pedagógicas mediante el diseño y aplicación de programas, proyectos y acciones.

Para conseguir los resultados de aprendizaje relacionadas con las dos CE mencionadas, el estudiante deberá desarrollar los contenidos en su vertiente teórica y aplicada que componen el programa de esta asignatura, junto con las Competencias Genéricas (CG) de la materia 8 a la que pertenece la asignatura de Evaluación de Programas, que le permitirán:

— Realizar estudios prospectivos y evaluativos sobre características, necesidades y demandas pedagógicas.

— Evaluar planes, programas, proyectos, acciones y recursos educativos y formativos desarrollados para atender a las necesidades y demandas previamente evaluadas.

Con ambos resultados se pretende contribuir al desarrollo del pedagogo como profesional competente en Evaluación de Programas.

De modo más concreto pretendemos que los estudiantes, tras cursar esta asignatura, consigan los siguientes resultados de aprendizaje, manifestados en su capacidad para:

1. Utilizar de modo preciso la terminología específica de evaluación de programas educativos.

2. Diseñar, planificar y realizar evaluaciones de programas educativos contextualizados, basados en la teoría de programas y de evaluación de programas, y elaborar sus informes.
3. Realizar evaluaciones de programas educativos seleccionando y aplicando los métodos de investigación de recogida y análisis de información adecuados a las características y propósitos pretendidos por el programa, teniendo en cuenta sus antecedentes, sus recursos y su entorno.
4. Elaborar propuestas de cambio y mejora de los programas evaluados a la luz de los resultados y de las nuevas necesidades detectadas.
5. Valorar y aplicar las normas o estándares del Joint Committee por las que se deben regir las evaluaciones de programas.
6. Interpretar y valorar informes de evaluaciones de programas de diferentes contextos y potenciar la utilización de sus resultados para la mejora de la calidad de la educación.

Los resultados de aprendizaje específicos se verían potenciados con el desarrollo de las siguientes competencias genéricas, que permitirán desarrollar en el alumno la capacidad para:

1. Comunicarse correctamente mediante el uso del lenguaje oral y escrito.
2. Utilizar las tecnologías de la comunicación y de la información, con fines de aprendizaje, para ampliar su formación.
3. Trabajar colaborativamente, de modo flexible, afrontando y colaborando en la solución de problemas, formulando posibles hipótesis para si solución, fundamentadas en la teorías y en la experiencia, de modo que aporte evidencias para el desarrollo de los programas y la atención a las necesidades.
4. Resolver problemas, analizando los hechos y las situaciones en las que reproduce.
5. Gestionar responsablemente su vida, conducir planes de desarrollo personal y profesional de futuro.
6. Aprender a lo largo de toda la vida para actualizar sus conocimientos y sus planes de desarrollo personal y profesional.

5. ORIENTACIONES AL ESTUDIO

Aconsejamos al estudiante que lea los contenidos de un modo crítico y reflexivo, interactuando con los argumentos procedentes de los representantes de los diferentes modelos de evaluación de programas. Que se acerque a la asignatura con una visión práctica, relacionando las concepciones teóricas de cada autor, los procesos que propone para realizar la evaluación y los métodos que utiliza, asociándolo todo a una realidad susceptible de ser mejorara y evaluada desde cada uno de los enfoques.

De esa manera el alumno podrá ir integrando los contenidos expositivos de la asignatura en un conjunto de elementos aplicables a entidades reales, lo que contribuirá a su formación teórica, procedimental, instrumental y resolutiva para realizar futuros trabajos evaluativos. La aplicación real o simulada de la evaluación de programas a un programa real, de modo individual y en grupo, supondrá la oportunidad de desarrollar actitudes de trabajo en grupo y colaborativo, necesarios para el desarrollo profesional que la vida laboral demanda.

En la práctica, el evaluador profesional puede elegir entre un variado rango de enfoques de evaluación, debe implicar a los interesados en la evaluación y escuchar sus puntos de vista, y como profesional de la evaluación debe utilizar las normas profesionales para asegurar y valorar la calidad de la elección de un determinado enfoque y concepción evaluativa de acuerdo con los propósitos de la evaluación, y así mismo estar sujeto a las evaluaciones de su evaluación.

La lectura comprensiva de los contenidos de este libro ayudarán al estudiante a valorar la importancia de la disciplina científica de la Evaluación de Programas para su formación como profesional de la Educación para contribuir a la mejora de los programas, de las organizaciones educativas y de la educación.

Aconsejamos que aproveche todos los recursos de apoyo al aprendizaje que la UNED pone a su servicio, desde el estudio de los materiales didácticos, la consulta de la bibliografía recomendada, la participación activa y colaborativa en los foros temáticos en el curso virtual de la asignatura, los documentos complementarios junto con la consulta específica a través de internet, los trabajos prácticos y las pruebas de evalua-

ción continua; la consulta telefónica al profesor de la Sede Central y al Profesor Tutor en los Centros Asociados, el uso del correo electrónico, la asistencia a las tutorías en los Centros Asociados y la utilización de las bibliotecas y las nuevas tecnologías para aprender.

Tema 1

La evaluación de programas educativos. Teoría de la evaluación de programas. Concepto y funciones de la evaluación de programas

1. Objetivos y resultados de aprendizaje.
2. Introducción.
3. Introducción a la evaluación de programas.
   - 3.1. La evaluación como disciplina de conocimiento.
   - 3.2. La evaluación de programas y la investigación evaluativa.
   - 3.3. Los métodos en la evaluación de programas.
   - 3.4. De los métodos a la teoría de la evaluación de programas.
   - 3.5. La naturaleza de la teoría de programas y de la evaluación de programas.
4. El concepto de evaluación de programas y los criterios y las referencias para la evaluación.
   - 4.1. Concepto de evaluación de programas.
   - 4.2. Los criterios y las referencias para la evaluación de programas.
5. Las funciones de la evaluación de programas.
   - 5.1. La evaluación formativa y sumativa.
   - 5.2. La evaluabilidad y su aplicación para la evaluación de los programas.
   - 5.3. La meta-evaluación y su función formativa y sumativa.
   - 5.4. La evaluación para el rendimiento de cuentas.

6. El objeto de la evaluación.
   6.1. La evaluación de los aprendizajes de los estudiantes.
   6.2. La evaluación de la calidad de los centros educativos y conceptos específicos.
      6.2.1. La evaluación de la gestión de la calidad de los centros educativos. Los principios de la gestión de la calidad total. Las organizaciones que aprenden.
      6.2.2. El aseguramiento de la calidad.
      6.2.3. Evaluación de programas y auditoría.
      6.2.4. Funciones del evaluador según su relación con la organización evaluada y el enfoque de evaluación utilizado.
7. Propósitos de la evaluación de programas.
8. Ejercicios de autoevaluación.
9. Referencias bibliográficas.

1. OBJETIVOS Y RESULTADOS DE APRENDIZAJE

Con este tema pretendemos que el estudiante comprenda y valore el concepto de evaluación de programas. De modo más concreto, esperamos que tras el estudio del tema el alumno consiga dominar los siguientes objetivos de aprendizaje:

— Comprender la importación de la teoría para la evaluación de programas.

— Analizar y valorar las diferentes definiciones o conceptos de «evaluación de programas» de los autores más representativos en este campo.

— Comprender y valorar los criterios y las referencias de las evaluaciones.

— Comprender y valorar las diferentes funciones de la evaluación de programas.

— Diferenciar la finalidad de la evaluación dependiendo de su objeto, ya sea la evaluación de programas, la evaluación del aprendizaje de los estudiantes o la evaluación de la calidad de los centros educativos.

2. INTRODUCCIÓN

La evaluación de programas es algo más que una metodología aplicada y, al igual que sucede en todas las disciplinas, requiere de sus propias bases teóricas que justifiquen y den sentido a lo que hace un evaluador, cómo lo hace y, sobre todo, para qué lo hace. El propósito de la evalua-

ción de un programa es informar a la acción, apoyar la toma de decisiones y aplicar el conocimiento para resolver problemas humanos y sociales; cuando este examen es realizado de manera sistemática y empírica mediante una cuidadosa recogida de datos y análisis rigurosos, se está realizando investigación evaluativa.

La investigación evaluativa utiliza las metodologías de investigación de las Ciencias Sociales. Es considerada una forma de investigación aplicada porque la evaluación de programas es, fundamentalmente, una actividad práctica centrada en la recogida de información relevante para un particular problema, programa o producto, currículo o una actividad en el aula, con la finalidad de informar la toma de decisiones de cambio y mejora de lo evaluado.

Las funciones educativas clásicas de la evaluación de programas, como son la evaluación formativa y sumativa, junto con las funciones psicológicas y administrativas de la evaluación, se han visto enriquecidas con otras nuevas que actualmente ha conquistado con el desarrollo de la gestión de la calidad de las organizaciones orientadas a la mejora continua, la innovación y la responsabilidad.

3. INTRODUCCIÓN A LA EVALUACIÓN DE PROGRAMAS

3.1. La evaluación como disciplina de conocimiento

La evaluación es el nombre de una disciplina autónoma con su propia clasificación en la Biblioteca del Congreso de los EEUU (Scriven, 1991:141). Una disciplina es un constructo intelectual organizado. La esencia de una disciplina es su trabajo y la clave para su existencia su concepto general y su metateoría. La disciplina de la Evaluación se centra en el estudio y aplicación de procedimientos para hacer evaluaciones objetivas y sistemáticas.

Dentro de esta disciplina autónoma, hay seis subáreas, o áreas semiautónomas, como son las disciplinas de evaluación de programas, de productos, de personal, de rendimiento, de proyectos y de políticas. Otros campos aplicados autónomos incluidos son la valoración tecnológica, la evaluación médico-psicológica y el control de la calidad.

En la emergencia de una nueva disciplina, de acuerdo con Scriven (1991:25), podemos distinguir los siguientes cuatro elementos:

1. Reconocimiento explícito de que algo nuevo está emergiendo y existe una definición de ese algo. Lograr su definición o esbozar sus limites conlleva intentos prescriptivos sobre los que descansan las normas y pautas de desarrollo de esa nueva disciplina. Es el primer paso hacia la metateoría, y ello se da en la Evaluación.

2. Identificación y desarrollo de una metodología apropiada, entendida como un grupo de procedimientos e instrumentos para generar resultados útiles en el nuevo campo. La metodología puede ser específica o puede abarcar un amplio rango o niveles de métodos generales y modelos de análisis, y esto último ocurre en la Evaluación.

3. Desarrollo del campo consistente en una amplia producción de investigaciones, de informes, de descripciones, de ilustraciones y de principios y teorías generales. Los elementos de su generalidad deben presentarse explícita o implícitamente al menos en forma de nuevos esquemas conceptuales, con la terminología específica asociada. Ello está representado por los múltiples modelos con un alto grado de elaboración en sus concepciones teóricas.

4. Desarrollo de una metateoría de la disciplina que proporcione un método de trabajo para la práctica. La metateoría es casi siempre descriptiva y prescriptiva derivando la última de las aportaciones de la primera. Sin una metateoría no puede identificarse el nuevo área de estudio como una nueva disciplina al no tener identificado y definido su territorio, no poder explicitar sus procedimientos, su metodología y sus límites. Y esto se ha podido extraer de la diversidad de modelos evaluativos y de su aplicación.

Los límites de la Evaluación como disciplina han sido definidos por la mayoría de los teóricos de la evaluación en términos que la consideran un área de investigación aplicada. La evaluación de programas utiliza métodos y técnicas de investigación que proceden de las ciencias sociales, y de otras disciplinas como el derecho, la lógica, la ética, y necesita del desarrollo de la teoría evaluativa.

La defensa de la Evaluación como disciplina descansa en sus características científicas en contraste con las evaluaciones subjetivas o no sistemáticas. La Evaluación es considerada como ciencia por estar comprometida con la producción de conocimiento, y no sólo conocimiento práctico; la evaluación, desde el punto de vista de la práctica, se interesa por conocer el mérito relativo de los programas, como tratamientos, como pueden ser los diferentes métodos de enseñanza, o las diferentes teorías de aprendizaje, o las nuevas concepciones en la práctica sanitaria, y los cambios que producen en el rendimiento, en el desarrollo personal y profesional, en la salud como consecuencia de la aplicación de dichos programas. La evaluación es esencial para distinguir la ciencia de la pseudociencia al evaluar la calidad de la evidencia científica, la calidad de los diseños de investigación, de los instrumentos de medida, de sus conclusiones, la calidad de los procedimientos en ingeniería, en las matemáticas, en los procesos judiciales. En el sentido de proceso intelectual, la evaluación forma parte de todas las disciplinas, tecnologías, manualidades y de pensamiento racional en general (Scriven, 1991:4). La Evaluación, por ello, es considerada una transdisciplina, definiendo este término, creado por Scriven, como una disciplina instrumental que sirve a otras disciplinas, como ocurre con la estadística o con la lógica. Del mismo modo que la disciplina de la lógica forma parte de la filosofía y tiene su propia metateoría y, además, es una transdisciplina por proporcionar servicios a otras disciplinas.

Scriven hace el siguiente análisis sobre la naturaleza de la Evaluación:

— La Evaluación no es solo una mera acumulación y síntesis de datos relevantes para la toma de decisiones. Digamos que éste es uno de los componentes de la evaluación y uno de sus propósitos.

— El otro componente, y previo al anterior, es el de las premisas o criterios evaluativos que verifican los valores relevantes y las normas elegidas para la realización de la evaluación.

La discusión metodológica acaecida en el último cuarto de siglo xx, desde diversos campos, ha sido fructífera para abrir a la evaluación su propio campo. Pero el paso fundamental para que la evaluación se constituya en una disciplina científica va más allá de la simple práctica evaluativa. Comienza identificando normas que gobiernen, que prescriban

cómo deben hacerse las evaluaciones (Martínez Mediano, 1998). Debemos ir más allá de la discusión metodológica y las explicaciones sobre la adecuación de una u otra metodología según el ámbito y propósito de la evaluación (meta-metodología), debemos desarrollar los procesos lógicos de las evaluaciones, así como el desarrollo de las teorías de programas y las evaluativas.

3.2. La evaluación de programas y la investigación evaluativa

Se han hecho muchos intentos para distinguir la evaluación de la investigación en términos de generalidad, replicabilidad y tipo de datos que utilizan ambos conceptos. La evaluación, cuando es realizada de un modo sistemático con miras a ampliar el conocimiento práctico y teórico es considerada investigación y cuando su finalidad es resolver un problema concreto, dentro de una situación particular, es considerada evaluación. En esta línea está el argumento de Stake y Denny (1969:374, en Worthen y Sanders, 1987:29):

> La principal diferencia entre evaluación e investigación es el grado por el cual los hallazgos son generalizables más allá de sus aplicaciones a un producto, programa o contexto particular. Casi siempre, los pasos adoptados por el investigador para obtener la generalización tienden a hacer sus investigaciones artificiales o irrelevantes a los ojos del práctico. El evaluador sacrifica la oportunidad de manipular y controlar, pero gana en relevancia para la situación inmediata. El investigador y el evaluador trabajan dentro del mismo paradigma investigativo pero desempeñan diferentes roles organizativos y atraen a diferentes audiencias.

El punto que separa la investigación evaluativa de la evaluación es el mayor rigor con que se realiza la primera, y su finalidad, la resolución de un problema real, de tal modo, que permita extrapolar la información y de las soluciones aportadas y aprender de ellas, para su aplicación en la solución de problemas similares.

Para Rossi (1982), y muchos otros especialistas, la evaluación siempre debe ser realizada con rigor y sistematicidad; tanto la evaluación como la investigación evaluativa suponen la aplicación sistemática de los proce-

dimientos de la investigación social para valorar la conceptualización y el diseño de los programas, su ejecución y su utilidad para la intervención social. Los procedimientos, en ambos casos, evaluación e investigación evaluativa, deben ser rigurosos para que los resultados y conclusiones que se puedan deducir sean fidedignos de la situación en la que se pretende resolver el problema. La investigación, y asimismo la evaluación, exige que el problema esté incardinado dentro de una teoría, que los procedimientos para resolverlos sean adecuados a las exigencias de la metodología de investigación científica y que sus resultados contribuyan al desarrollo de dicha teoría.

3.3. Los métodos en la evaluación de programas

La utilización del método científico en la investigación educativa y social supone un intento sistemático dirigido a cuestionar y/o resolver problemas a través de la recogida y análisis de datos primarios con el propósito de describir, explicar, generalizar y predecir, siendo las características propias del método científico la repetibilidad, la estabilidad y la comunicabilidad. El enfoque de evaluación orientada por el método se vio claramente apoyada por el hecho de que la mayoría de las perspectivas de evaluación han estado orientadas principalmente por el método en los comienzos del desarrollo de la evaluación de programas. Por ejemplo, para Campbell y Stanley (1963) y Cook y Campbell (1969), defensores del paradigma experimental, una evaluación estaba mejor realizada si seguía la rigurosidad de los diseños experimentales.

Maslow, en su publicación de 1966, *The Psychology of Science (La Psicología de la Ciencia),* distinguía entre la ciencia mecanicista (la defendida por Kerlinger) y la ciencia humanista. En opinión de Maslow, la concepción mecanicista de la ciencia defendida por los positivistas no es incorrecta pero sí demasiado estrecha y limitada (Maslow, 1966: 5). Una concepción, más inclusiva podría ser aquella que recoge lo idiográfico, lo experiencial, el taoísmo, lo comprensivo, lo holístico, lo personal, lo transcendental, lo final (ibíd.: 63), en definitiva, dice, si hay una norma principal sobre la ciencia es la aceptación del reconocimiento y la descripción de todas las realidades, de todo lo que existe (ibíd.: 12). La ciencia, desde este punto de vista, tiene como finalidad comprender y trans-

formar para mejorar las condiciones humanas; es idiográfica, es decir, se interesa por lo individual y lo particular sin pretender llegar al establecimiento de leyes generales.

Es interesante decir que la investigación naturalista fue mencionada por primera vez como una alternativa a los métodos convencionales en el campo de la investigación evaluativa. El desarrollo parece haber comenzado con Parlet y Halmilton en 1972, en Gran Bretaña, con su idea de evaluación iluminativa, así como con R. Stake cuando definió la evaluación respondiente o sensible, en 1975, en Estados Unidos de América. Edmon Guba, en 1978, en una monografía, también lo recoge, pero su mayor expansión fue hecha por Michael Q. Patton con su obra *Qualitative evaluation methods (Métodos cualitativos en evaluación)*. Los defensores del enfoque naturalista argumentaban que los métodos cualitativos servían mejor a la evaluación. E, incluso, investigadores naturalistas como Parlett y Hamilton (1978) y Guba y Lincoln (1981) negaban cualquier mérito de los métodos experimentales o cuasi-experimentales aplicados a la evaluación. Las evaluaciones orientadas por los métodos cobraron fuerza por ese largo debate entre los métodos de investigación cuantitativos y cualitativos más adaptables a la evaluación.

Cronbach, en 1982, argumentaba que el uso exclusivo de los métodos cuantitativos basándose en su rigurosidad, lleva a que una evaluación sea rígida y estrecha en su alcance. Estaba interesado en ampliar la evaluación de programas a los procesos en los contextos y en las organizaciones para facilitar una toma de decisiones plural; su enfoque utiliza tanto métodos cuantitativos como cualitativos y defiende que la flexibilidad de los métodos cualitativos sirve mejor a las necesidades de una evaluación. Sostiene que la flexibilidad de los métodos cualitativos es útil para lograr la generalización de los resultados de la investigación y sirven también para aportar información a los patrocinadores de los programas. La generalización, según Cronbach, puede hacerse para subpoblaciones próximas y similares a las estudiadas, mediante un informe detallado, vívido y coherente con las necesidades de las distintas audiencias.

El uso de una metodología u otra para la evaluación de un programa dependerá de los propósitos de la evaluación, del estado de madurez del programa, de los recursos disponibles, de los intereses y necesidades de los patrocinadores y de los evaluadores y del medio político y organizati-

vo del programa. Así, cuando un programa está desarrollándose, el estudiar su eficacia mediante un diseño experimental carece de sentido puesto que pasaría de puntillas sobre los procesos y los contextos. Del mismo modo, utilizar los métodos cualitativos, en exclusividad, puede no ser lo más adecuado para valorar la eficacia de un programa de amplio alcance, elaborado para aplicarlo a toda la población de un país.

El propio Cook (1985) ha adoptado una perspectiva postexperimental, de «multiplicidad crítica» ya que defiende que la perspectiva del paradigma experimental debe transformarse en otra más comprehensiva que pueda adaptarse mejor a los aspectos políticos y organizativos de la evaluación de programas. Cook, junto a sus colaboradores, Shadish y otros, afirma que los evaluadores deberían planificar sus evaluaciones previendo el uso de múltiples metodologías, investigar múltiples aspectos de los programas y considerar el punto de vista de múltiples participantes. Cook dice que esta nueva perspectiva tiene ventajas sobre el paradigma experimental y considera que reduce las posibilidades de interpretación errónea, a la vez que proporciona una información más comprensiva, global e integradora para los procesos políticos, además de favorecer una evaluación más racional y consciente de sus valores.

Sin un esfuerzo hacia una fundamentación de por qué es positiva la complementariedad metodológica, la defensa del uso de múltiples metodologías en la evaluación puede no ser apoyada. Se necesita una conceptualización y un esfuerzo teórico para sistematizar e integrar los factores conceptuales y los métodos de investigación para la evaluación de programas. La teoría de programas debe proporcionar orientaciones para identificar qué temas son los más importantes en una evaluación, qué método o métodos son los más relevantes, y sugerir cómo aplicarlos para alcanzar los objetivos propuestos en la evaluación.

Este enfoque de complementariedad metodológica deriva del pragmatismo. Aplica métodos de investigación tanto cuantitativos como cualitativos para atender a las necesidades y propósitos para cada fase de la investigación. Defienden la complementariedad y uso de metodologías mixtas, centrando su atención en el problema de la investigación. Los métodos cualitativos y los cuantitativos utilizados en la investigación de modo complementario supone una nueva concepción de investigación, y es que la complejidad de los fenómenos educativos sólo podrá ser capta-

da y explicada mediante la utilización de enfoques flexibles, mixtos o complementarios que los avances en la metodología de investigación y su previsible desarrollo nos ofrecen.

La evaluación de programas, o investigación evaluativa, abarca diversos grupos de actividades según se trate del estudio de la conceptualización y diseño de las intervenciones de programas, la valoración de la puesta en marcha del programa, o la valoración de su utilidad y de su impacto, y, aunque no siempre es posible evaluar todo este grupo de aspectos relacionados con los programas, dependiendo de los recursos y tiempo, hay ocasiones en las que sí es necesario realizar una evaluación integral, una evaluación global (comprehensive evaluation), que es el enfoque evaluativo defendido por Rossi y Freeman (1989).

Los modelos de evaluación de programas que defienden la complementariedad metodológica aportan criterios suficientes que muestran la conveniencia de esta complementariedad metodológica dentro de sus propios modelos.

3.4. De los métodos a la teoría de la evaluación de programas

Una definición del término teoría es aquél que dice que la teoría proporciona un marco de referencia para ayudar a las personas a comprender su mundo y a actuar en él. La teoría es crucial para la investigación; no sólo proporciona orientaciones para analizar un fenómeno sino que proporciona un esquema para comprender el significado de los resultados de la investigación.

La tendencia en evaluación ha consistido en que las evaluaciones estén orientadas por los métodos, y, aunque esto ha supuesto importantes contribuciones para el desarrollo de la evaluación de programas, de otro lado, ha ocasionado que se descuidara el desarrollo del marco conceptual que guiara la práctica evaluativa. Esto se explica porque la Evaluación de programas no se veía como una disciplina de conocimiento, sino como una modalidad de investigación aplicada dentro de la Educación y de las Ciencias Sociales. De hecho, la acción y la práctica preceden al desarrollo de la teoría en cualquier disciplina, y a la Evaluación de programas le ha ocurrido lo mismo.

Una evaluación no puede realizarse sin limitaciones de tiempo y de recursos, y no puede abarcar todas las opciones, y ello supone llegar a acuerdos que requieren no sólo de unas metodologías sino de más de un grupo de principios y de teorías que guíen el proceso de evaluación de programas. Shadish y cols. (1986) dicen que el reto más importante que tiene planteada esa multiplicidad crítica es el desarrollo de la teoría evaluativa. La teoría evaluativa tiene que ver con la teoría de los programas y con los procedimientos para guiar la evaluación desde el análisis de las necesidades de los programas en sus contextos hasta el análisis de su planificación, de su aplicación y de sus resultados.

Entre las ventajas de la realización de evaluaciones guiadas por la teoría de programas está la de que la teoría permite comprobar si los resultados del programa se deben a como se ejecuta el programa, a la conceptualización propia del mismo o a ambos. La teoría de programas ayuda a clarificar si un programa está siendo desarrollado siguiendo los criterios explicitados en el mismo y en las condiciones adecuadas. La teoría de programas puede usarse para especificar los efectos intermedios de un programa que podría llegar a evidenciar los aspectos positivos y negativos del mismo antes de que los resultados finales puedan manifestarse, lo que proporcionaría posibilidades de mejora del programa antes de que se haya concluido el programa.

La incorporación de la teoría de programas en una evaluación fortalece la información de los evaluadores para clarificar las metas del programa, para observar los procesos y para atender las necesidades de los patrocinadores y de los participantes y beneficiarios, y para orientarles en la realización de las diferentes decisiones alternativas de mejora. Como dice Bickman (1987) la evaluación es, a menudo, considerada como una ciencia práctica, pero es ambas cosas, una práctica y una ciencia que requiere de su propia teoría.

3.5. La naturaleza de la teoría de programas y de la evaluación de programas

Un programa de intervención educativa o de intervención social supone un esfuerzo intencionado y organizado para intervenir en un proceso

con el propósito de resolver un problema y proporcionar un servicio. La estructuración de los esfuerzos organizados adecuadamente y dirigidos a la consecución de las metas deseadas, supone que el programa opera bajo una teoría, aunque, a veces, esta teoría esté implícita.

La teoría explícita proporciona una guía general para la elaboración del programa y explica como el programa organiza su funcionamiento, aunque puede verse influida por los prejuicios. La teoría, como los prejuicios, en la evaluación de programas, debe estar explícita.

El significado del término teoría más aceptado es el que la define como un grupo de concepciones interrelacionadas, principios o proposiciones para explicar o guiar acciones sociales.

Weiss (1972) sugería que la utilización de los resultados de una evaluación podría verse reforzada si se incluyeran estos elementos básicos:

1. Un análisis de las premisas teóricas del programa.
2. Una especificación de los procesos del programa para informar acerca de los vínculos entre los objetivos y las metas del programa.
3. Un análisis de las realizaciones del programa y de sus resultados y
4. Un informe detallado de cuáles son las alternativas más efectivas que pueden ayudar a mejorar el programa y cómo realizarlas.

Una de las razones del movimiento para el desarrollo de *la teoría de la evaluación de programas* se relaciona con la necesidad de comprender cómo se realizan los programas, considerados como tratamientos para solucionar un problema. Los estudios centrados en la realización de los programas y en la aplicación de los mismos aportan evidencias de que su aplicación es algo muy complicado dentro del ámbito humano. Las dificultades que se han revelado en la puesta en marcha y desarrollo de los programas han señalado la necesidad de que los promotores de un programa presten atención a estos aspectos para mejorar la implementación del programa. Es importante, por ello que, dentro del desarrollo del marco teórico de la evaluación de programas, se integre el desarrollo de los programas en la acción, como parte importante de la actividad evaluativa. Se trataría con ello de evitar lo que a veces ocurre cuando se evalúan los programas, que quienes realizan dichas evaluaciones tienden a pro-

ceder más según sus puntos de vista y su preparación metodológica que en función de las necesidades reales de ese programa y esto no siempre es consistente con los objetivos de quienes elaboran o diseñan un programa adaptado a unas necesidades para resolverlas.

Desde el punto de vista conceptual los evaluadores no sólo deben informar de la realización de sus evaluaciones sino que deben trabajar con los patrocinadores del programa para conocer y mejorar los modos de implementación de la realización de los programas, y posteriormente, deben colaborar para mejorar el programa y su aplicación. Deben informar sobre los cambios que se requiere hacer en el programa, qué criterios los fundamentan, qué estrategias se deben planificar para lograrlos, quiénes deben realizar los cambios y qué consecuencias, previsibles y no previsibles, producirán las modificaciones para la mejora (De la Orden, 1996).

Chen (1990), para argumentar su *Teoría de evaluación de programas*, toma una idea de Weber. El desarrollo de la evaluación de programas y su emergencia como disciplina se ha visto beneficiada por el desarrollo de la sociedad moderna, racionalista y contable.

Sin embargo, Weber advertía que es importante ver la diferencia entre dos tipos de racionalidad: la racionalidad formal y la racionalidad sustantiva, o los valores. La racionalidad formal tiene que ver con la eficiencia y la eficacia en la consecución de una meta específica en un tiempo dado, mientras que la racionalidad de valor o sustantiva se refiere a los propósitos sustantivos y fines a largo plazo de los individuos, de los grupos y de las sociedades. Cuando la racionalidad formal se dirige a una explicación parcial, estrecha, de las actividades sociales, los propósitos generales y las metas a largo plazo se pierden de vista. El argumento de Weber es relevante para la conceptualización de la evaluación de programas. Cuando su conceptualización es estrecha y simplemente mide las metas conseguidas, el énfasis se pone en la racionalidad formal, los evaluadores se interesan por la eficiencia y por la eficacia del tratamiento, descuidando algo tan esencial como la cuestión sobre la adecuación de las metas a las necesidades y las consecuencias intencionadas y no intencionadas a medio y largo plazo. Para evitar tales problemas se necesita ampliar el marco conceptual de la evaluación de programas, de modo que esté más implicada en la racionalidad sustantiva o de valor,

que proporcione una información más rica y real a los propósitos de los programas y a sus implicaciones en intereses sociales más amplios.

En una teoría de evaluación de programas, los criterios seleccionados para juzgar la realización de un programa claramente implican juicios de valor. Temas claves relacionados con cómo diseñar un tratamiento, o cómo llevarlo a cabo, también implican juicios de valor. Podemos decir que la teoría, los valores, las propias preferencias, e incluso los prejuicios, deben hacerse explícitos. No cabe duda de que cualquier opción estará influida por la teoría y por los prejuicios. (Shadish, Cook & Leviton, 1991:31 y ss.).

La teoría de la evaluación de programas es definida por Chen (1990) como *una especificación de lo que se debe hacer para conseguir las metas deseadas, qué otros impactos o logros importantes pueden ser también anticipados, y cómo estas metas y estos impactos pueden generarse*.

La primera parte de esta definición se relaciona con la teoría prescriptiva, también llamada normativa, ¿qué se debe hacer? y la segunda parte se relaciona con la teoría descriptiva, también llamada causativa, en lo que subyace a los mecanismos causales que relacionan los tratamientos del programa, los procesos de realización, los procesos de implementación y las metas. Chen prefiere utilizar la terminología de *teoría normativa* en lugar de teoría prescriptiva y *teoría causativa* en vez de teoría descriptiva. Ambas subteorías forman parte de la teoría de la evaluación de programas.

La evaluación desde la teoría normativa requiere de la consistencia entre la estructura teórica del programa y la estructura de funcionamiento del programa para ser valorada. Esto puede realizarse en un estadio temprano del desarrollo del programa que proporcionará información para diagnosticar pronto los problemas, a tiempo, y realizar acciones para tratar de que estos problemas se subsanen, fortaleciendo la estructura del programa y/o los procesos de realización.

Cuando un programa de intervención —o tratamiento— es construido y ejecutado adecuadamente, la credibilidad del programa se potencia. No obstante, esto no lleva implícito que el programa sea eficaz. La evaluación del programa, desde la teoría normativa es insuficiente por sí misma para comprender si el programa consigue las metas para las cuales ha sido diseñado o para conocer las consecuencias no intencionadas

del programa. Para comprender la eficacia del programa, la teoría descriptiva, explicativa o causativa es importante y necesaria.

Una evaluación que se rige por la teoría causativa proporciona información sobre el impacto generado por el programa y sobre cómo dichos impactos fueron generados. El valor de un programa es muy difícil de juzgar sin tener información sobre factores de contexto o de intervención que ayuden a conocer por qué un programa tiene éxito o fracasa.

La especificación de la teoría causativa en una evaluación es especialmente útil para la mejora futura del programa. Cuando la teoría causativa es proyectada en detalle la evaluación puede conocer las debilidades de los mecanismos causales que subyacen en el programa, identificar los factores contextuales e intervinientes que dificultan o facilitan los procesos del programa y sus resultados, y sugerir posibles estrategias para ocuparse de mejorar sus programas e introducir los cambios necesarios para ello. Ambas teorías, normativa y causativa, son necesarias y constituyen las dos subteorías de la teoría de la evaluación de programas. (Martínez Mediano, 1998).

La teoría de programas orienta la teoría de la evaluación de programas.

4. EL CONCEPTO DE EVALUACIÓN DE PROGRAMAS Y LOS CRITERIOS Y LAS REFERENCIAS PARA LA EVALUACIÓN

4.1. Concepto de evaluación de programas

Ralph Tyler (1950:69), considerado el fundador de la evaluación aplicada a los programas educativos, concebía la evaluación como «El proceso para determinar en que medida los objetivos educativos han sido realizados».

Los criterios que orientan la evaluación de los programas son los objetivos que los programas pretenden conseguir. Los programas educativos son creados para atender a unas necesidades rigurosamente evaluadas y definidas, siendo ese el objetivo que pretenden conseguir. Los programas se diseñan fundamentados en teorías que han demostrado su valía para la acción, con una estructura, estrategias, actividades y temporalización orientadas a la resolución de las necesidades definidas, y los

objetivos de los programas son los criterios que guían la evaluación del programa, en su diseño, puesta en acción, realización y logros; los objetivos son los criterios frente a los cuales se contrastan los resultados conseguidos por el programa tras su aplicación, si bien la evaluación de programas no solo se centra en los objetivos conseguidos sino que considera todo lo que contribuye al logro de esos resultados.

En la raíz de la «Evaluación» está el término «*valor*», indicando que la evaluación esencialmente implica la realización de *juicios de valor*. Pero estos juicios de valor deben estar fundamentados en principios defendibles y criterios que guíen la evaluación y determinen el valor de lo evaluado frente a ellos.

Los evaluadores deben elegir los criterios de valor apropiados para juzgar lo que evalúan de acuerdo con los objetivos del programa para responder a las necesidades. Y deben señalar el mérito en lo conseguido, fijando un *estándar de realización* para determinar si lo que se está evaluando ha trabajado y rendido de acuerdo con los criterios que perseguía, los objetivos, y si ha conseguido unos resultados de calidad de acuerdo con los estándares fijados.

La evaluación debe comprobar que los programas están diseñados al servicio de los objetivos que pretende conseguir, —los criterios— y que estos objetivos, tras la aplicación del programa, se consiguen con suficiencia, con calidad —*el mérito*—. Y esta comprobación debe ser hecha mediante la *recogida de información de modo sistemático, rigurso, válido y fiable*, sobre las necesidades a las que pretende responder, sobre el programa, sus metas y sus objetivos en coherencia con las necesidades, sobre la planificación de su puesta en acción teniendo en cuenta los recursos personales y materiales, sobre sus acciones aplicativas y sobre sus resultados inmediatos y su impacto, siguiendo las características propias de la investigación.

El término evaluación en lengua española se ha utilizado fundamentalmente referido a la evaluación del rendimiento de los estudiantes pero debido al creciente desarrollo del concepto de evaluación aplicado a los programas, los currículos, al personal y los centros educativos, se hace necesario especificar el ámbito de la evaluación al que nos referimos. En inglés existen diversos términos utilizados para evaluar, tales como: *evaluation*, *as-*

sessment, *appraisal*. Y entre los evaluadores que han estudiado estos términos, como Choppin, (1990) o Scriven (1991), han llegado al acuerdo de utilizar el término *evaluation* de modo genérico, aplicado a la disciplina de Evaluación, con sus diferentes subcampos, como son la «evaluación de programas», la «evaluación del personal», la «evaluación de centros educativos», la «evaluación de sistemas educativos» o la «evaluación de políticas», y apoyan el uso del término *assessment* para la «evaluación del rendimiento de los estudiantes», o de las personas. El término *appraisal,* que se traduce como «tasación», o «valoración», suele utilizarse con connotaciones de aplicación a «valores de mercado» o referido a «objetos».

Para Scriven (1980:18), la evaluación «es el proceso de determinar el valor y la calidad de algo, siendo las evaluaciones el resultado de dicho proceso». La evaluación es la valoración sistemática del valor y mérito de lo que se evalúa; es una actividad profesional que incluye procedimientos sistemáticos para la recogida de información que debe culminar en juicios fundamentados y con recomendaciones objetivas.

Cronbach *et al.* (1980:14) definen la evaluación de programas como un

> ... examen sistemático de los acontecimientos que ocurren durante el proceso de aplicación de un programa y las consecuencias que se derivan del mismo, con el fin de ayudar a su mejora y a la de otros programas con similares propósitos y características.

Para Stufflebeam y Shinkfield (1987:166)

> ... la evaluación es un proceso mediante el cual se proporciona información útil y descriptiva para la toma de decisiones. En un sentido más amplio, la evaluación consiste en un proceso en tres etapas para identificar, obtener y proporcionar información acerca de un programa, valorado en sus metas, en su planificación, en su realización y en su impacto, con el propósito de guiar la toma de decisiones, proporcionar información y contribuir a su comprensión, con el criterio de su valor (respuesta a las necesidades) y su mérito (calidad).

Pérez Juste (2000) define la evaluación como

> ... el proceso sistemático de recogida de información rigurosa (valiosa, válida y fiable) orientado a valorar la calidad y los logros de un programa, como base para la posterior toma de decisiones de mejora tanto del programa como del personal implicado.

Y para la autora de esta obra, la evaluación de un programa elaborado para atender a unas necesidades supone la recogida de información sistemática sobre la calidad del propio programa, en sus fundamentos, estructura y coherencia interna; sobre su planificación para su puesta en acción, en el contexto de una organización, considerando los profesionales que lo aplican y los medios materiales; sobre el programa en la acción, en sus resultados inmediatos y en sus resultados en el medio y largo plazo, o su impacto, con el objetivo de comprobar y valorar el grado y calidad con el que el programa ha sido elaborado, planificado, implementado y si consigue los resultados de acuerdo con las metas pretendidas y si ha logrado otros valores. De este modo, la evaluación proporciona información basada en hechos acerca de las potencialidades del programa en su contexto para recomendar su continuidad y expansión, o sus necesidades de cambio y mejora en aquellos aspectos señalados por la evaluación. La evaluación de programas, aplicada en todo su alcance, constituye una estrategia de aprendizaje continuo para la mejora de los programas, de las personas, de la institución en la que se desarrollan y de la sociedad a la que sirve.

El potencial de la evaluación para la mejora de la calidad de los programas y de las organizaciones es de una gran trascendencia al ayudar a identificar los aspectos positivos y negativos con los que operan los programas y las organizaciones, junto con los diferentes elementos que contribuyen al logro de sus objetivos, sobre los que hay que incidir para introducir cambios para mejorarlos y contribuir así a la mejora de los programas y de sus resultados en las personas, en las organizaciones y en la sociedad.

4.2. Los criterios y las referencias para la evaluación de programas

Los criterios de la evaluación son los fines de la educación. Dicho de otro modo, los criterios son los objetivos de los programas, sus dimensiones. La evaluación debe asegurar que sean de calidad, que sean coherentes con las necesidades y que se cumplan. Evaluar implica valorar de acuerdo con unos criterios, y determinar el mérito y la calidad de lo que consigue el programa. La comparación entre los criterios y la aplicación del programa, durante su proceso y tras su aplicación para valorar sus

resultados, nos permitirá atribuir valor sobre el propio programa, su aplicación y sus resultados, aportando información sobre lo evaluado, teniendo en cuenta las condiciones del contexto, y todo ello nos permitirá elaborar recomendaciones. Scriven denomina estos elementos como el «mérito, valor y significación de lo evaluado».

Lo que la evaluación dice que es adecuado y positivo, es lo que orienta a mantener los fines de la educación, los objetivos, o cambiarlos.

Los principales criterios que se utilizan para la evaluación de la calidad de los programas son los siguientes:

— Pertinencia: se define como la medida en la cual un programa responde a una o más necesidades concretas de la población o grupo al que atiende.

— Eficacia: es la medida en la cual existe prueba de que los objetivos establecidos por el programa se han conseguido. La evaluación de la eficacia tiene como finalidad comprobar el grado en que los objetivos pretendidos por un programa han sido conseguidos.

— Eficiencia: relación entre los recursos utilizados y los resultados conseguidos.

— Efectividad: Es la medida en la cual un programa ha alcanzado una serie de efectos.

— Satisfacción al final del programa, que se añade a la conseguida a lo largo del programa en relación con las expectativas sobre el programa.

— Impacto, que son los efectos posteriores a la terminación del programa originados por el programa. Los criterios para valorar la eficacia y el impacto son los objetivos declarados que pretende conseguir el programa para resolver determinadas necesidades.

Las referencias o los estándares de la evaluación guardan relación con los criterios y definen cómo deben cumplirse los criterios para poder emitir un juicio positivo sobre la intervención del programa. Los estándares deben ser realistas, señalando una graduación en su consecución en función de diversos elementos contextuales. Son tres las referencias que se utilizan: la referencia criterial, la normativa y la idiosincrática.

— Si la referencia se toma en relación con el punto de partida se dice que es idiosincrática, y se centra en el progreso conseguido. El punto de partida pueden ser los resultados de la evaluación última, que se considera la línea base frente a la cual se comparan los resultados que el programa consigue en el centro, pudiendo decir que los resultados conseguidos con las mejoras introducidas en el programa y su aplicación, mejoran los resultados en relación con los que el programa conseguía en el período anterior.

— Si es en relación con otros programas, o con otros centros que utilizan el mismo programa, hablaremos de referencia normativa, en relación con una distribución continua de resultados, y se centra en conocer si los resultados han sido superiores, inferiores o iguales que los conseguidos por otros programas u otros centros.

— Si lo hace en relación con un dominio, se llama referencia criterial, y se compara lo que el programa ha conseguido con lo que pretendía conseguir a un nivel absoluto, centrado en el logro del dominio.

5. LAS FUNCIONES DE LA EVALUACIÓN DE PROGRAMAS

Son varias las funciones que cumple la evaluación dependiendo del propósito de la misma, del tipo de datos y del momento en el que se realiza. En el momento de la planificación de una evaluación es importante determinar las funciones específicas que cumplirá la evaluación, debiendo informar de ellas a los implicados en la evaluación.

5.1. La evaluación formativa y sumativa

Scriven (1967) formuló dos importantes funciones de la evaluación: la formativa, para la mejora del programa mientras se desarrolla, y la sumativa, valorando la eficacia del programa en su totalidad, cuando ha terminado.

La *evaluación formativa* es realizada desde el inicio del programa y durante su desarrollo con el fin de recoger información útil y a tiempo de los efectos del programa y de las necesidades con el fin de introducir

cambios para mejorar el programa y sus resultados. Informa a la evaluación sumativa sobre el funcionamiento del programa. La evaluación formativa es fundamental para la mejora continua de lo que se está evaluando —un programa, un material didáctico, el aprendizaje de un alumno—, y además, debe informar sobre el funcionamiento del programa, al servicio de la evaluación sumativa.

La *evaluación sumativa* es realizada una vez que ha terminado el programa, o en determinados momentos clave, y su objetivo principal es recoger información sobre los resultados finales y sobre el valor y mérito del programa. Sirve al rendimiento de cuentas y a la mejora de los programas en su totalidad. La evaluación sumativa, además de informar sobre los resultados del programa, cumple funciones formativas sobre el programa concluido, para orientar a la comunidad para mejorarlo.

5.2. La evaluabilidad y su aplicación para la evaluación de los programas

El concepto de evaluabilidad es definido como la medida en que un programa puede ser evaluado (Wholey, 1994). La evaluabilidad es una actividad previa a la evaluación y determina si el programa obedece a unas necesidades y si está fundamentado en una teoría para la acción, su aplicación, en coherencia con los fines para los que dice ser su razón, y si se realizan.

Todo programa debe pretender conseguir unos objetivos para atender a unas necesidades evaluadas y definidas previamente, y debe explicar cómo va a actuar para lograr dichos objetivos mediante un programa con unos fundamentos teóricos específicos. De ese modo, si se trata de un programa educativo, éste debe explicar las teorías de aprendizaje en las que se fundamenta de modo explícito, o el evaluador deberá hacerla explícitas, y sus actividades, para el logro de los objetivos, formando parte de la estructura y coherencia interna del programa. Asimismo, debe tener definida la población a la que el programa se dirige, así como las expectativas de logro del grupo al que atiende.

El programa educativo debe tener definidos los objetivos que el programa afirma pretender conseguir, la coherencia de la teoría que subya-

ce a dicho programa y las propuestas y acciones que debe realizar para conseguir los resultados pretendidos y comprobar sus efectos en los sujetos a los que se dirige. Y para su adecuada aplicación debe contar una estructura organizativa coherente que le aporte cobertura, las personas responsables de su aplicación y unos recursos materiales que hagan posible la aplicación y desarrollo del programa. Y debe describir unas estrategias, actividades y procedimientos para el logro de los objetivos que se propone conseguir. El programa debe definir unos indicadores de ejecución, de modo que informe sobre cómo debe ser implementado el programa y facilitar comprobar si se están logrando los resultados pretendidos durante el proceso que conducen al logro de los resultados a conseguir. Debe concretar las acciones a realizar mediante los procesos, con una duración definida en el tiempo, señalando su comienzo y su final, y debe especificar los procedimientos que va a utilizar para el seguimiento de lo que el programa consigue en momentos clave durante la aplicación del programa y para la evaluación de los resultados finales pretendidos por el programa y los efectos en los sujetos a los que se dirige y otros.

Si no existiera programa, o si existiendo no se aplica de acuerdo a como está previsto, si se desarrolla teniendo poco que ver con el plan previsto, si la evaluación no puede realizarse con el mínimo rigor y colaboración de los interesados, entonces podría ser desaconsejable realizar la evaluación. Por lo tanto, la evaluabilidad atiende a la evaluación de las necesidades, los objetivos en coherencia con las mismas, el programa para resolverlas y su implementación, junto con los interés de los implicados en el programa y el uso posible de sus hallazgos.

Si el programa no estuviera bien definido, su aplicación tuviera poco que ver con el programa o los fines de la evaluación no estuvieran claramente determinados en coherencias con las necesidades y los implicados en el programa no participan y colaboran en la evaluación, podría hacer no aconsejable la realización de la evaluación del programa. Y la evaluabilidad del programa, basándose en la revisión del mismo y de cómo se está aplicando, valorará si procede realizar su evaluación.

También se interesa la evaluabilidad por el tipo de información y datos que debe y pueden recogerse en la evaluación del programa, los mo-

mentos en los que debe recogerse y la finalidad de la evaluación, de acuerdo con los intereses de los grupos implicados en la evaluación y el uso de los resultados de la evaluación.

5.3. La meta-evaluación y su función formativa y sumativa

La meta-evaluación, término acuñado por Scriven (1967) y también utilizado por Stufflebeam (1981), y con posterioridad a ellos por todos los especialistas en evaluación, se define, por analogía con el término evaluación, como una investigación sistemática encaminada a emitir un juicio valorativo acerca de la calidad y los méritos de una evaluación.

Atendiendo a la defensa de Scriven en relación con la meta-evaluación, ésta debe utilizarse como imperativo profesional, con el fin de asegurar la calidad y honradez, precisión y veracidad del trabajo de evaluación realizado reflejado en sus informes. La posible influencia de naturaleza política ha de ser considerada en cualquier análisis meta-evaluativo que pretenda comprender la evaluación en toda su complejidad.

Para Stufflebeam (1987:52)

> El objetivo de la meta-evaluación es asegurar la calidad de los servicios evaluativos, evitar, o enfrentarse a las prácticas ilegales o a los servicios que no son de interés público, señalar el camino para el perfeccionamiento de la profesión y promover una mayor comprensión de la actividad evaluativa.

La meta-evaluación no puede verse solo como la mera revisión de la calidad técnica con la que se hace una evaluación, sino que se constituye en una función estratégica que debe actuar como salvaguarda tanto de los procedimientos seguidos en la evaluación, cuanto de sus fines y de sus resultados.

En la tercera revisión de las normas para la evaluación de programas educativos (Yarbrough, et al., 2011), se enfatiza la necesidad de mejorar y asegurar la responsabilidad de las evaluaciones mediante la meta-evaluación sistemática, tanto interna como externa, con propósitos formativos y de mejora de la evaluación y con propósitos de responsabilidad o de rendición de cuentas.

5.4. La evaluación para el rendimiento de cuentas

El incremento del énfasis dado en la evaluación de programas al rendimiento de cuentas se debe, fundamentalmente, al compromiso de los responsables de los programas de informar sobre sus resultados a quienes tienen derecho a saber, los implicados en su aplicación, los que lo reciben, los padres, la administración, los accionistas y la sociedad. El rendimiento de cuentas en evaluación tiene que ver con el ejercicio de la responsabilidad como complemento al ejercicio de la autonomía para desarrollar programas educativos en instituciones apoyadas económicamente con fondos estatales o fondos externos. Tiene que ver con la transparencia sobre cómo se desarrollan los programas educativos y los resultados que consiguen. Consiste, básicamente, en informar a quienes tienen derecho a saber acerca de cómo ha funcionado una institución para desarrollar lo que tenía encomendado.

Los centros educativos están apoyados económicamente mediante los presupuestos del Estado, o de otros benefactores o accionistas, y por tanto deben responder del uso de la confianza depositada en ellos. Así, por ejemplo, si se quiere optar a una determinada ayuda o subvención para desarrollar un programa, o un proyecto —o si dedica parte de los presupuestos de educación a la implantación de nuevos programas, o de reformas educativas—, lleva como contraprestación informar sobre los resultados que dicho proyecto o programa obtiene a lo largo de los diversos momentos de la aplicación y cuando finaliza. Si los resultados conseguidos no son los previstos, esto podría tener consecuencias para pedir responsabilidades y para la obtención de aprobaciones de nuevos apoyos, nuevas ayudas o subvenciones.

Cuando se presenta un proyecto de desarrollo educativo, social, un proyecto de investigación, y se concursa para conseguir financiación, tiene que incluirse un plan de evaluación que muestre los objetivos del proyecto, el uso de los recursos, las actividades y los resultados en el corto, medio y largo plazo y el plan de evaluación desarrollado para la comprobación del diseño, aplicación y resultados del proyecto, fundamentado, incluyendo los momentos, los objetos y los instrumentos de evaluación, y, después debe mostrar los resultados que ha conseguido en relación con los pretendidos, y las propuestas de mejora. Es la manera de informar

del uso de los fondos públicos en las investigaciones, y el modo de aportar transparencia sobre lo realizado, los logros alcanzados y la justificación de nuevas propuestas.

Así, diversas organizaciones, a cambio de recibir subvenciones del gobierno, tienen que rendir cuentas del logro de las metas que afirman pretender conseguir, debiendo validar, acreditar o certificar la consecución de esas metas a través de auditorías realizadas por agencias externas. La publicación de los informes de resultados de organismos gestionados con fondos públicos, o privados, aporta transparencia sobre las metas, procedimientos y resultados, y permite a los ciudadanos, en coherencia, elegir los servicios que le ofrece una u otra institución.

Esta necesidad de rendir cuentas —*accountability*—, o de garantizar o asegurar la calidad de lo que se hace —*quality assurance*—, ha creado la necesidad de conocer y utilizar los conceptos y funciones de la evaluación de programas, sus metodologías y procedimientos en las auditorías. En consecuencia, los auditores han reconocido la necesidad de conocer métodos diferentes a los de la auditoría, que les permitiera recoger información sobre temas relacionados con las metas de los programas y sus actividades, tareas éstas relacionadas con la actividad del evaluador de programas (Chelimsky 1997: 6). Por otro lado, los expertos en evaluación de la calidad de las instituciones suelen utilizar el término «gestión» para la medida y valoración del funcionamiento y rendimiento institucional, y de la satisfacción de los beneficiarios directos e indirectos con ese rendimiento.

6. EL OBJETO DE LA EVALUACIÓN

Si bien este libro está dirigido a la evaluación de programas, hay muchos otros objetos posibles de evaluación. Desde la evaluación de los estudiantes, la evaluación del personal, la evaluación de las organizaciones, la evaluación de políticas educativas, sociales, o de cualquier otro tipo, la evaluación de equipos informáticos, de servicios, de teorías, etc.

6.1. La evaluación de los aprendizajes de los estudiantes

La evaluación del estudiante es el área de la evaluación utilizado y desarrollado con anterioridad al desarrollo teórico y de aplicación de la evaluación de programas, lo que le ha permitido prestar una gran ayuda a la evaluación de programas. La evaluación del aprendizaje de los estudiantes es una de las funciones propias realizadas por el profesor

En la evaluación del rendimiento de los alumnos, se plantea el problema del criterio o la norma a utilizar para interpretar adecuadamente el progreso en el aprendizaje que logra el alumno tras un determinado período instructivo. El criterio es el referente frente al cual se compara el rendimiento alcanzado con unos niveles previamente fijados y nos permite emitir un juicio evaluativo sobre el aprendizaje logrado por el estudiante de suspenso, aprobado, notable, sobresaliente o matrícula de honor.

En la evaluación referida a norma, el aprendizaje logrado por un alumno se determina comparándolo con el obtenido por un grupo de referencia pudiendo ser el grupo de compañeros de su aula, o de su curso. Si la puntuación que alcanza es igual o superior a la puntuación media del grupo de referencia, el alumno aprobará, descansando el criterio de aprobar o suspender en datos ajenos al rendimiento concreto del sujeto. Esto representa una medida relativa del aprendizaje, porque el énfasis no recae sobre el aprendizaje real conseguido por el alumno, si no en la posición relativa de éste con un grupo normativo de referencia. Y podremos decir que el aprendizaje conseguido por el alumno, o su rendimiento, se sitúa en relación con una distribución normal, en torno a la media aritmética, por encima de le media, por debajo de la media.

Los tests psicométricos como son los tests de inteligencia general, de aptitudes o de personalidad, utilizan criterios normativos de comparación de la puntuación obtenida por un sujeto en relación con su grupo normativo de referencia y clasifican la posición del sujeto con respecto a la media del grupo o en relación con los cuartiles o los percentiles.

La evaluación de los aprendizajes con referencia a criterio representan el aprendizaje real de un estudiante con referencia a un dominio, entendiendo el término tal como un conjunto de conocimientos, habilidades, destrezas y competencias que deben ser dominados por el alumno para

considerar que ha superado un programa de estudio y diremos si ha aprendido todos los contenidos y destrezas, y con qué calidad, lo que nos permitirá calificar el aprendizaje del alumno como aprobado, notable, sobresaliente, o suspenso. El criterio es lo ideal; la consideración de si un alumno debe o no promocionar, aprobar, superar el programa de estudios, debe fijarse mediante un punto de corte, por ejemplo, en el 5 en una escala de 1 a 10 (Popham, 1983; Martínez Mediano, García Llamas & Del Rio, 1997).

Si la evaluación del aprendizaje de los estudiantes tiene como referencia el progreso personal alcanzado por los estudiantes de modo individual en función de su programa de estudios y sus capacidades y posibilidades personales, hablaremos de evaluación idiosincrática o personalizada. La valoración de los resultados en el aprendizaje se realiza mediante la comparación de los resultados obtenidos por el alumno al finalizar el programa con sus conocimientos iniciales, con la línea base del estudiante en ese programa, y con ello nos referiremos al progreso personal del alumno y diremos que mejora progresivamente, notablemente, o insuficientemente, y se valorará como un rendimiento insuficiente, suficiente o satisfactorio.

6.2. La evaluación de la calidad de los centros educativos y conceptos específicos

La evaluación de programas, y sus modelos, persiguen evaluar un programa diseñado y aplicado para atender unas necesidades previamente diagnosticadas que es necesario subsanar. Un programa educativo, dentro de un currículo, busca desarrollar unas determinadas habilidades, conocimientos, actitudes, competencias. Los programas buscan lograr unos resultados mediante su aplicación en organizaciones educativas, con unos recursos económicos y equipos de profesores formados para aplicar dichos programas a determinados grupos de alumnos sobre los que se pretenden conseguir unos cambios; y para conseguirlo toda la institución debe estar al servicio de una adecuada aplicación de los proyectos y programas educativos.

La evaluación de programas no se limita a evaluar únicamente el programa, sino que debe tener en cuenta el contexto en el que se desarrolla,

las necesidades que motivan los programas, la organización, las personas que los aplican. La actividad evaluativa está integrada en las organizaciones educativas. La evaluación debe centrarse en comprobar el funcionamiento del propio programa, junto con todos los elementos necesarios para que el programa consiga los cambios que pretende, en las actitudes, en los conocimientos y en la formación de los que aplican el programa.

La evaluación de programas no puede estar aislada de la mejora de la calidad de las organizaciones educativas y de la calidad de la educación en general. La evaluación de un programa no se puede aislar de su contexto, de sus responsables, de sus principales actores, sin tener en cuenta las metas de la organización, sus recursos, sus beneficiarios, la legalidad vigente, la transparencia, la información fundamentada para su utilización para mejorar el programa en sucesivas aplicaciones, al servicio de los programas y de sus beneficiarios, alumnos, profesores y sociedad.

La evaluación de programas y de los centros corresponde al profesional de la evaluación de programas y de centros. Puede ser realizada por el profesor o por un equipo de profesores integrado en una unidad de calidad, con conocimientos especializados sobre evaluación, cumpliendo funciones de evaluación interna de la calidad de los programas en su contexto, para la mejora de los programas y de sus resultados y de la organización del centro, y también puede y debe hacerse, además, por un evaluador profesional externo al centro para ofrecer objetividad y transparencia a los resultados de la evaluación.

6.2.1. La evaluación de la gestión de la calidad de los centros educativos. Los principios de la gestión de la calidad total. Las organizaciones que aprenden

La evaluación de programas guarda relación con la evaluación de los sistemas de gestión de la calidad y, del mismo modo, el trabajo del evaluador guarda relación con el trabajo del auditor de los sistemas de gestión de la calidad y de los evaluadores de los modelos de excelencia. La gestión de la calidad se orienta siempre hacia el logro de lo que la institución pretende conseguir, recogiendo evidencias sobre cómo lo hace y lo que consigue, orientada a la mejora continua, en el convencimiento de que nada es tan

bueno que no pueda ser mejorado. Por ello, una finalidad de la gestión de calidad es la mejora continua de las organizaciones hacia la excelencia.

Las organizaciones educativas participan de este interés por la excelencia; sin perder de vista su propia realidad, tratan de incorporar los principios de la calidad total mediante la implantación de modelos de gestión de la calidad adaptados a las organizaciones educativas. (Maldonado Rivera & Martínez Mediano, 2015:20; Martínez-Mediano & Arribas Díaz, 2016). Con la implantación de los sistemas de gestión para el aseguramiento de la calidad por las organizaciones, se asume la necesidad de evaluar lo que en dichas organizaciones se realiza.

Tanto en los modelos de evaluación de la calidad de los programas y de las organizaciones como en los modelos para la gestión de la calidad, concurren los objetivos de evaluación para la mejora y para el rendimiento de cuentas, como aseguramiento de la calidad. La evaluación de la calidad como un proceso de mejora continua se fundamenta en el ciclo de Deming o PDCA *(plan, do, check, act),* en español planifica, haz, comprueba y actúa en consecuencia (PHCA) (Deming, 1989), de gestión de la calidad total, que culmina en un informe de resultados.

La función del rendimiento de cuentas, como aseguramiento externo de la calidad, ha mostrado su influencia en el desarrollo de la cultura por la calidad en las instituciones de educación superior, dado que los resultados de estos procesos evaluativos no terminan en un mero informe, sino que incluyen sugerencias de mejora que deben ponerse en la acción y que son objeto de su seguimiento y evaluación posterior, basados en la confianza mutua y en la transparencia (Maldonado Rivera & Martínez Mediano, 2015:20-21). Ante una evaluación, la mayor parte de la sociedad espera que se le ofrezca las claves de mejora del sistema.

En general, los modelos de gestión de la calidad contemplan la «responsabilidad social» como uno de los conceptos fundamentales y se diseñan en función del contexto histórico social en el cual se circunscriben y se desarrollan.

En todos los modelos subyacen los principios básicos de la calidad total, a saber: una dirección o liderazgo para establecer políticas y coordinar esfuerzos, una dirección estratégica con objetivos anuales, un equipo de personas motivadas y formadas para implementar la estrate-

gia, unos recursos necesarios, unos procesos determinados con metodología propia orientada hacia la satisfacción de las expectativas de los clientes, de marketing y venta del producto o servicio; y unos beneficios o resultados de rendimiento.

Las organizaciones inteligentes o las organizaciones que aprenden utilizan la evaluación para mejorar, aplican programas de aseguramiento de la calidad, de garantía de la calidad, de gestión de la calidad, a modo manuales de calidad, donde se elabora un proyecto de centro, de lo que quiere conseguir, cómo lo quiere conseguir y cómo va a comprobar lo que consigue, evalúan sus innovaciones para informar y mejorar en aquello que sea necesario en relación con la organización, sus profesionales, sus programas, sus servicios y resultados, atendiendo las necesidades de sus beneficiarios y de la sociedad. Como puede observarse evaluar un programa en su contexto, y todo lo que lo hace posible, es una tarea compleja y la evaluación de programas ofrece procedimientos fundamentados, enfoques de investigación mixta, complementaria, y sugiere la necesidad de que la evaluación sea realizada en equipo.

6.2.2. El aseguramiento de la calidad

El concepto de aseguramiento de la calidad, o garantía de la calidad, se basa en la elaboración, por parte de una organización, de una propuesta con el compromiso de aplicarla bien a la primera, cumpliendo con los objetivos propuestos y de acuerdo con los criterios de calidad previamente definidos, comprometiéndose a aportar información sobre su cumplimiento.

El aseguramiento de la calidad, o garantía de la calidad, es un proceso sistemático diseñado para identificar, analizar y eliminar variaciones en los proyectos, procesos y resultados. Asume unas normas o estándares métricos, así como el análisis de las mejores prácticas en su ámbito. Pone especial énfasis en la concepción del proyecto o del programa y en el trabajo por procesos sujetos a evaluación, con el objetivo puesto en la mejora continua. Las necesidades y expectativas de los beneficiarios o clientes, tanto internos como externos, son los árbitros últimos de la valoración de la calidad y del éxito del programa.

El aseguramiento de la calidad se orienta a la «satisfacción de las necesidades y expectativas del cliente». El aseguramiento de la calidad está adoptado en todos los ámbitos. Las organizaciones deben adelantar su Sistema de Aseguramiento de la Calidad, o de garantía de la calidad , que informe que ese proyecto o programa esta planificado y será realizado, ajustado a principios claros de calidad, de modo que garanticen que se va a realizar y a obtener los resultados pretendidos, incluidos los gastos presupuestados. Los agentes externos valorarán el propio Sistema de Gestión de la Calidad (SGC) de la institución, su aplicación y sus resultados. De ese modo se rinde cuentas, se informa de un modo claro, justificado, argumentado y transparente de lo que consigue una organización en relación con lo que afirma pretende conseguir.

La «evaluación de programas» se relaciona con la expresión «aseguramiento de la calidad». El término «aseguramiento de la calidad» es un término prestado de los ámbitos empresariales o industriales, dentro del movimiento de la Calidad que se utiliza también en Educación. El aseguramiento de la calidad es considerado un concepto más amplio, que contiene e interpreta al de evaluación de programas, desde una perspectiva de la calidad (Maldonado Rivera & Martínez Mediano, 2015, 2016).

Estos enfoques han sido adaptados para su aplicación en los ámbitos de las organizaciones educativas, como por ejemplo en la educación superior para la evaluación de su calidad, utilizando modelos sistémicos, que incorporan estrategias para la mejora continua propias de las filosofías de la Calidad Total.

6.2.3. Evaluación de programas y auditoría

Centramos la comparación entre la auditoría de realización —no la auditoría financiera— y la evaluación de una organización. El auditor y el evaluador se plantean diferentes tipos de preguntas. El auditor está interesado en saber si lo que se ha hecho está conforme con lo que se debía haber hecho; busca respuestas a preguntas del tipo ¿qué? El evaluador trata de comprender que es lo que produce ciertos efectos ya sean deseados o no deseados, explícitos o no explícitos; busca respuestas a preguntas de tipo ¿por qué?

El auditor puede estar comprometido con defender la transparencia y el derecho a saber, y el evaluador considera que una de sus mayores contribuciones es reformular lo que se está haciendo, y anima a los responsables a buscar soluciones, e incluso a ver los problemas de diferente modo. Ambos, auditor y evaluador, pueden estar comprometidos con incrementar el control, el derecho a saber de la colectividad sobre los programas y las organizaciones financiadas con fondos públicos o privados, pero desde distintos enfoques; el auditor utiliza métodos retrospectivos, contrastando el manual de funcionamiento —manual de calidad— y los objetivos pretendidos con los resultados, con el fin de detectar errores, incumplimientos, no conformidades, en relación con lo que se dice que se va a hacer en el manual de calidad; el evaluador utiliza metodologías retrospectivas y prospectivas, con el fin de detectar a qué se han podido deber los incumplimientos y proponer nuevas vías para conseguir las metas pretendidas junto con nuevas metas valoradas de interés dentro de los proyectos educativos del centro las programaciones de aula y los proyectos transversales y específicos, para la mejora de la formación de los estudiantes.

Las actividades de auditoría son supervisadas por las organizaciones auditoras reconocidas; los auditores reciben sobre su formación inicial una formación orientada a las políticas prioritarias en cada período de tiempo. Los evaluadores deben adaptarse a las características de los programas y de las organizaciones a evaluar y completar su información y formación, si fuera necesario, de acuerdo con las características propias de lo que se va a evaluar.

Los auditores requieren una formación que va más allá de las habilidades contables, siendo una formación específica con unas características ampliamente aceptadas por las agencias nacionales (Entidad Nacional para la Acreditación y Certificación —ENAC—, en España) e internacionales de auditoría, si bien puntualmente sea necesaria la constitución de equipos con consultores especializados en función de las actividades que desarrollen las organizaciones y los programas que se auditen. También los evaluadores deben constituirse en equipos de expertos para manejar los problemas con maestría, frecuentemente relacionados con las metodologías de investigación y con los fundamentos de los programas educativos.

Los auditores, cada vez más, constatan que deben conseguir colaboración y apoyo de aquellos a los que auditan, y esto también es válido

para los evaluadores. Los diseños de evaluación y de auditoría son discutidos y modificados con los miembros de las organizaciones a auditar o evaluar y reciben ayuda explícita de algún miembro de la institución. Los auditores tienen el acceso a la empresa a auditar, normalmente por ley, y asimismo acceden a toda la documentación; su continuidad en la actividad auditora, año tras año, también suele estar garantizada. Los evaluadores deben negociar su acceso a la organización, y no siempre se les facilita toda la información.

Los auditores planifican sus auditorías de un modo pragmático para su realización en unas pocas jornadas de trabajo y utilizan herramientas de recogida de información funcionales. Los evaluadores diseñan sus evaluaciones siguiendo modelos de la investigación social, recurren a la teoría sustantiva y tras la recogida de información mediante diversidad de técnicas elaboran informes que pretenden sirvan a la propia organización así como al desarrollo de teorías. Debemos señalar que hay dos tipos de auditoría fundamentales orientados a la certificación de la implantación de los Sistemas de Gestión de la Calidad de las Normas ISO (International Standardization Organization): La auditoría interna, realizada por un equipo interno del centro preparado en ISO 9000, y la auditoría externa, realizada por un equipo externo acreditado por las entidades acreditadoras competentes para realizar dicho trabajo.

6.2.4. Funciones del evaluador según su relación con la organización evaluada y el enfoque de evaluación utilizado

Si la evaluación es realizada por los miembros de una organización, con fines diagnósticos y de mejora, se habla de *evaluación interna,* pudiendo nombrarse como *autoevaluación;* si la evaluación es realizada por expertos externos a la institución, con fines de rendimiento de cuentas, se habla de *evaluación externa*.

La *autoevaluación,* referida a las instituciones, es la herramienta fundamental que utilizan los modelos de evaluación global de las organizaciones, como son los modelos de excelencia, entendida como un examen global y sistemático de los proyectos, actividades y resultados de una organización que se compara con criterios facilitadores y de resultados del

modelo de excelencia. Mediante el proceso de autoevaluación se pretende conseguir una comprensión detallada de la organización, buscando realizar un diagnóstico sobre su situación actual, señalando los puntos fuertes y las áreas de mejora. A partir de dicho diagnóstico considerado como línea base, se elaborarán planes de mejora concretos, realistas y evaluables cuya aplicación y resultados serán evaluados, introduciéndose en la organización una dinámica de mejora continúa, o lo que es lo mismo, la *gestión de la calidad*.

Cuando se inicia un proyecto de evaluación de una institución se habla de *evaluación institucional,* y presupone las dos funciones de la evaluación, *la evaluación interna, o autoevaluación,* para aportar información para introducir mejoras inmediatas, y *la evaluación externa,* para informar, como rendimiento de cuentas ante quienes tienen derecho a saber sobre el funcionamiento de la institución, o para optar a un reconocimiento como son los premios de Excelencia del modelo Europeo de Excelencia de la Fundación Europea para la Gestión de la Calidad de la European Foundation for Quality Management (EFQM) (Martínez Mediano & Riopérez, 2005).

Desde el enfoque de evaluación institucional, se cubre el doble objetivo de conocer el funcionamiento de la institución, con un primer objetivo de *evaluación diagnóstica*, para introducir los ajustes necesarios en la conceptualización del programa, en la organización, en la gestión de los recursos materiales, la formación del personal y los medios necesarios para realizar las mejoras, incluidos los económicos. El segundo objetivo sería el de informar a la sociedad sobre cómo trabaja la institución para conseguir las metas que se ha marcado y los logros que consigue, a modo de rendimiento de cuentas, fundamentada esta información en la evaluación interna, o autoevaluación, y en la evaluación externa, realizada por evaluadores externos, lo que garantiza su credibilidad.

7. PROPÓSITOS DE LA EVALUACIÓN DE PROGRAMAS

La evaluación tiene, entre uno de sus objetivos, el estudio de las políticas públicas y la adecuación de sus presupuestos, debatidos en el seno de las instituciones políticas, como son los parlamentos. El impulso ha-

cia la consolidación de la evaluación y de su institucionalización, procede, fundamentalmente, de los poderes públicos, en forma de financiación dada por el poder legislativo para la evaluación, que la han visto como un elemento de legitimación de sus políticas y del uso de los fondos públicos para su financiación a través de los informes de expertos sobre la adecuación de las mismas y de su aplicación. Por ello, entre otras razones, las evaluaciones se realizan con el propósito de informar e influir en el proceso de toma de decisiones políticas.

En síntesis, son tres los propósitos de la evaluación de programas:

— Emitir juicios sobre el valor de un programa, fundamentados, basándose en un análisis comprensivo del programa, de cómo se aplica y los resultados que consigue.

— Asistir al responsable de la toma de decisiones, para la mejora, aportando un *feedback* del propio programa y su aplicación, considerando los recursos personales, materiales y las propias estrategias y actividades para la aplicación del programa.

— Cumplir una función social, y política, asumiendo responsabilidades, rindiendo cuentas *(accountability)* tanto en relación con los aspectos económicos como sociales y políticos.

La evaluación se ha constituido en una disciplina necesaria para conocer el funcionamiento de los programas de cualquier ámbito de la actividad humana: político, económico, social, sanitario, educativo, de las instituciones, de las organizaciones y de los grupos que los desarrollan. El propósito de la investigación evaluativa es informar a la acción, apoyar la toma de decisiones y aplicar el conocimiento para resolver problemas humanos y sociales.

EJERCICIOS DE AUTOEVALUACIÓN

1. Valore críticamente la importancia de la teoría de programas para la evaluación.
2. Valore las distintas definiciones de evaluación y señale elementos comunes y diferenciales entre ellos.

3. Describa una actividad de un programa educativo para desarrollar un objetivo del programa y explique como la evaluaría aplicando una definición concreta de evaluación de programas.
4. Funciones de la evaluación formativa y la evaluación sumativa y su contribución a la mejora de los programas y de las organizaciones educativas.
5. Concepto de evaluabilidad y explique como se aplicaría en relación con un programa educativo que pretende evaluar.
6. Concepto de meta-evaluación y sus funciones en la evaluación de programas.
7. Concepto de rendimiento de cuentas.
8. Elementos comunes entre los criterios para la evaluación de programas y para la evaluación de los estudiantes. Reflexione sobre su relación.
9. Haga un glosario de los términos claves del tema, y continúe con esta actividad en cada uno de los temas del programa de estudio de esta asignatura.

REFERENCIAS BIBLIOGRÁFICAS

Arribas Díaz, J. A. & Martínez Mediano, C. (2015), «El programa AUDIT de la ANECA y las normas internacionales ISO 9000. Análisis comparativo», *Educación XX1*, 18, 2, 375-395. DOI: 10.5944/educXX1.13938

Bickman, L. (Ed.) (1987). *Using programm theroy in evaluation*. San Francisco: Jossey-Bass.

Chelimsky, E. (1997). «The political environment of evaluation and what it means for the development of the field. Preface». En E. Chelinsky y W. R. Shadish (Eds.) *Evaluation for the 21st century*. Thousand Oaks, CA, London, New Delhi: Sage.

Chen, H-T. (1990). *Theory-driven evaluation*. Newbury Park: Sage.

Choppin, B. H. (1990). «Evaluation, assestment and measurement». En H. J. Walberg y G. D. Haertel (1990): *The international encyclopedia of educational evaluation*. Oxford, NY: Pergamon Press. (7-8).

Cook, T. D. (1985). «Postpositivist critical multiplism». *Social Science and Social Policy,* 21-62.

Cook, T. D. y Campbell, D. T. (1969). *Quasi-experimentation: Designs and analysis issues for field settings.* Chicago, ILL.: Rand-McNally.

Cook, T. D. y Reichardt, Ch. S. (1986). *Métodos cualitativos y cuantitativos en investigación educativa.* Madrid: Morata.

Cronbach, L. J. (1982). *Designing evaluations of educational and social programs.* San Francisco: Jossey-Bass.

Cronbach, L. J. *et al.* (1980). *Toward reform of program evaluation.* San Francisco, CA: Jossey-Bass.

Cronbach, L. J. & Suppes, P. (Eds.) (1969). *Research for tomorrow's schools: disciplined inquiry for education.* New York: Macmillan.

De la Orden, A. (1996). *Evaluación, innovación y calidad educativa. XI Congreso Nacional de Pedagogía: Innovación pedagógica y políticas educativas.* San Sebastián: Sociedad Española de Pedagogía y Universidad del País Vasco.

EFQM (2013). *Introducción al Modelo EFQM de Excelencia.* Asociación Europea de Gestión de la Calidad (EFQM) y Club de Excelencia en la Gestión.

Guba, E. G. (1978). *Toward a methodology of naturalistic inquiry in educational evaluation*. Los Angeles, CA: Center for the Study of Evaluation.

Guba, E. G. y Lincoln, I. S. (1981). *Effective evaluation.* San Francisco, Washington, London: Jossey Bass.

ISO 9000:2015. *Quality Management System. Fundamentals and Vocabulary.* International Standard Organization. www.iso.org

Kerlinger, F.N. (1973). *Foundation of Behavioral Research.* 2nd Edition. Holt Rinehart and Winston.

Maldonado Rivera, J. J. & Martínez Mediano, C. (2015). «Gestión de la calidad de postgrados a distancia en América Latina y el Caribe». *Fundamentos y análisis comparativo.* Série-Estudos, 40, (15-34).

— (2016). «Análisis de los modelos de evaluación de posgrados a distancia en América Latina y el Caribe». *Revista de pedagogía,* Bordón, 68, 3, (73-90).

Martínez Mediano, C. (1998). «La teoría de la Evaluación de Programas». *Educación XX1,* ISSN: 1139-613X. 73-91. DOI: http://dx.doi.org/10.5944/educxx1.1.1.398 http://revistas.uned.es/index.php/educacionXX1/article/view/398

Martínez Mediano, C., & Riopérez Losada, N. (2005). «El Modelo de Excelencia de la EFQM y su aplicación para la mejora de la calidad de los Centros Educativos». *Educación XX1,* 8, 35-65. DOI: http://dx.doi.org/10.5944/educxx1.8.0.342 http://revistas.uned.es/index.php/educacionXX1/article/view/342

Martínez Mediano, C. & Arribas Díaz, J. A. (2016). *Valoración de la Aplicación de los Sistemas de Gestión de la Calidad ISO 9001 en los Centros Educativos a través de un Cuestionario.* Meta: Avaliação, 8, 23, (326-358).

Martínez Mediano, C., García Llamas J. L. & Del Río Sadornil, D. (1997): «Evaluación de los aprendizajes de los alumnos» (173-206). En Martínez Mediano, C. (Coord.). *Encuentros en la Facultad de Educación sobre Evaluación.* Madrid. UNED, Serie Actas y Congresos (37153PB01).

Maslow, A. H. (1966). *The psychology of science.* New York: Harper & Row.

Parlett, M. R. y Hamilton, D. (1978). «Evaluation as illumnation: A new approach to the study of innovative programmes». *Occasional paper,* n.º 9, Centre for Research in the Educational Sciences. University of Edinburgh. 1972.

Patton, M. Q. (1990). *Qualitative evaluation and research methods.* Newbury Park, CA: Sage.

Perez Juste, R. (2000), (Coord.). «Evaluación de programas educativos». En *Revista de Investigación Educativa*. N.º monográfico, 18, 2.

Popham, W. J. (1983). *Evaluación basada en criterios*. Madrid: Magisterio Español.

Rossi, P. H. (1982). «Some dissenting comments on Stake's review». En E. R. House, S. Mathison, J. A. Pearsol y H. Preskill (eds). *Evaluation studies review annual* (vol.7) Beverly Hills, CA: Sage.

Scriven, M. (1967). «The methodology of evaluation». En R. W. Tyler, R. W. Gagné y M. Scriven (eds.), (1968): *Perspectives of curriculum evaluation.* Chicago, Ill. Rand McNally.

— (1980). *The logic of evaluation.* Inverness, CA: Edgepress.

— (1991). *Evaluation Thesaurus.* Newbury Park, CA: Sage.

Shadish, W. R. Cook, T. D. y Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice.* Newbury Park, CA: Sage.

Stake, R. E. (1975) (Ed.), *Evaluating the arts in education: An responsive approach* (13-31), Columbus, OH: Merrill.

Stake, R. E. y Denny, T. (1969). «Needed concepts and techniques for utilizing more fully the potential of evaluation». En R. W. Tyler (Ed.): *Sixty-eight year book of the National Society for Study of Education. Part II: Educational Evaluation: new roles, new means*. Chicago, Ill.: University of Chicago Press.

Stufflebeam, D. L. & Shinkfield, A. J. (1987): *Evaluación sistemática. Guía teórica y práctica.* Madrid: Paidós, MEC.

Tyler, R. W. (1950). *Basic principles of curriculum and instruction.* Chicago: University of Chicago Press.

Weiss, C. H. (1975). *Investigación evaluativa: Métodos para determinar la eficiencia de los programas en acción.* México: Trillas.

Wholey, J. S. (1994). «Assessing the feasibility and likely usefulness of evaluation». En J. S. Wholey, H. P. Hatry, y K. E. Newcomer (Eds.). *Handbook of Practical Program Evaluation.* San Francisco: Jossey Bass.

Worthen, B. R & Sanders, J. R. (1987). *Educational evaluation: Alternative approaches and practical guidelines*. New York, London: Longman.

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.

Tema 2

Michael Scriven. La evaluación como la ciencia de la valoración. La evaluación orientada al consumidor. La lista de las claves de control para la evaluación y la meta-evaluación

1. Objetivos y resultados de aprendizaje.
2. Introducción.
3. La teoría de la evaluación de Scriven.
   - 3.1. La evaluación como la ciencia del valor.
   - 3.2. Las necesidades sociales y la ciencia de la valoración.
   - 3.3. La evaluación libre de costes.
   - 3.4. La evitación de los prejuicios.
   - 3.5. La evaluación libre de metas
   - 3.6. La evaluación formativa y sumativa.
   - 3.7. La meta-evaluación.
   - 3.8. La lógica de la evaluación y la práctica evaluativa.
4. La evaluación orientada al consumidor.
5. La lista de las claves de control para la evaluación y la meta-evaluación.
6. Tabla resumen del enfoque de evaluación de programas de Scriven.

Ejercicios de autoevaluación.

Referencias bibliográficas.

1. OBJETIVOS Y RESULTADOS DE APRENDIZAJE

Con el estudio de este tema se pretende que el estudiante comprenda y valore las aportaciones de Scriven a la evaluación de programas. De manera más concreta se espera que el alumno sea capaz de:

— Comprender el concepto de evaluación de programas de Scriven.

— Valorar y aplicar los conceptos de evaluación formativa y sumativa en la evaluación de programas.

— Valorar la concepción de Scriven de la evaluación como la ciencia de la valoración y su aplicación a la evaluación de programas en sus contextos.

— Comprender la lógica de la evaluación para la realización de evaluaciones.

— Comprender y ser capaz de aplicar procedimientos para la evitación de los prejuicios en la evaluación.

— Aplicar la lista de claves de control para la evaluación y la meta-evaluación.

— Comprender, valorar y aplicar la meta-evaluación formativa y sumativa.

2. INTRODUCCIÓN

Michael Scriven, británico de nacimiento pero con nacionalidad estadounidense, estudió Matemáticas y se doctoró en Filosofía en Oxford.

Ha sido profesor en prestigiosas universidades de Estados Unidos, Reino Unido, Australia y Nueva Zelanda. Ha sido presidente de la American Educational Research Association (AERA) y de la Asociación Americana de Evaluación. En la actualidad está vinculado a la Western Michigan University, de Estados Unidos.

Scriven es uno de los primeros teóricos de la evaluación moderna y ha influido más que ningún otro en cómo se habla en evaluación. Su artículo *The methodology of evaluation* (1967) ha sido uno de los más citados en evaluación educativa. Los términos *evaluación sumativa* y *evaluación formativa* fueron introducidos por él en 1967; la *evaluación libre de metas* y la meta-evaluación son también términos suyos, además del término inglés *evaluand* (el objeto a evaluar) y la *evaluación libre de costes*. En su obra titulada *Evaluación Thesaurus* (1991) recoge la terminología específica de la evaluación de programas.

Scriven ha sido el gran impulsor para el desarrollo de la Evaluación como disciplina de conocimiento. Una disciplina es un constructo intelectual organizado, siendo su esencia su práctica, junto con su concepto general y su meta-teoría.

La evaluación debe *proporcionar información para los que toman decisiones,* pero también debe *determinar el mérito y el valor de lo que se evalúa*. Scriven es el más persistente defensor de esta posición; para él la evaluación es *la ciencia de la valoración*. La clave para realizar una buena evaluación, de acuerdo con Scriven, consiste en determinar las premisas o criterios que orientan las metas de los programas y su realización, y la evaluación, basándose en esos criterios debe verificar que dichos criterios y valores relevantes se consiguen.

El desarrollo de las listas de las claves de control para la evaluación y la meta-evaluación también es una contribución de Scriven, habiendo mostrado su utilidad e importancia para la realización de evaluaciones de programas.

3. LA TEORÍA DE LA EVALUACIÓN DE SCRIVEN

Para Scriven, la evaluación es la valoración sistemática del valor y mérito de las cosas, de lo que se evalúa. La evaluación es una actividad profesional que incluye procedimientos sistemáticos para la recogida de información, que debe culminar con juicios bien fundamentados, y con recomendaciones objetivas.

3.1. La evaluación como la ciencia del valor

La mayoría de los evaluadores reconocen que la evaluación asigna mérito o valor pero raramente profundizan en ello. Scriven es el más persistente defensor de esta posición. Para él, la evaluación es *la ciencia de la valoración*. Critica a otros evaluadores por definir la evaluación como *aquello que proporciona información para los que toman decisiones*. «La evaluación es lo que es, la determinación del mérito o el valor y para lo que sea usada es otro asunto» (1980:7, en Shadish, Cook y Leviton, 1991:74). Ese otro asunto depende del ámbito en el que se realice la evaluación. Si es en el educativo, es claro que la evaluación se utiliza para tomar decisiones basadas en la información acerca de cómo funciona el programa, su mérito y el valor de cómo lo hace, y en las síntesis valorativas o juicios valorativos, teniendo en cuenta las situaciones en las que se desarrollan los programas, para aconsejar si conviene o no modificar un programa educativo, en qué aspectos y cómo; por ello, para lo que es usada es importante. Pero, ciertamente, la evaluación es un proceso y un resultado, y el juicio sobre el objeto evaluado es una de las funciones inexcusables a realizar por el evaluador.

Cuestiona las evaluaciones basadas en la tradición tylerina que consideran que la evaluación debe determinar si los objetivos han sido alcanzados, considerándolas demasiado limitadas para prever otros posibles efectos. Los evaluadores deben juzgar las metas en sí de los programas, pero no limitarse por ellas. Deben identificar todos los resultados que consigue un programa, no solo los relacionados con los objetivos propuestos, y valorar si han dado respuesta a las necesidades y a las expectativas de los beneficiarios del programa y si añaden valor.

Scriven cree que los valores pueden ser investigados y justificados empíricamente, y, en este sentido defiende la objetividad de los criterios de valoración para la realización de evaluaciones. Considera que los evaluadores deben ser capaces de llegar a juicios de valor justificables, más allá que medir resultados o concluir que las metas han sido alcanzadas. Opina que los argumentos para mantener a la ciencia libre de valores son, en general, inadecuados porque el trabajo de los científicos tiene implicaciones de valor, lo quieran o no, lo pretendan o no, de modo que «la investigación evaluativa debe producir, como conclusión, un juicio de valor» (1974a:4). Argumenta que las afirmaciones de valor son similares a otros constructos científicos: «El valor, el mérito, la calidad son constructos procedentes de variables observables, lo mismo que las aptitudes, el rendimiento, la motivación o la ansiedad» (1986:39).

Los valores son un caso especial de constructos científicos:

> El valor de los programas, el mérito de los profesores, la calidad de los resultados son, simplemente, constructos teóricos que pueden ser indefinidamente encontrados dentro de las implicaciones factuales, enlazados en la red de nuestros constructos, necesidades y medios utilizados como cualquier otro constructo en el mundo de las acciones de la práctica, o en el mundo del intelecto. No es posible aceptar el papel de los constructos empíricos en las ciencias y mantener cualquier argumento en contra del papel de los constructos de valor (Scriven, 1983:41).

Scriven se declara neopositivista reconociendo que ningún científico observa la realidad sin distorsionarla, la observa a través de sus filtros perceptivos y sus prejuicios. Por ello, la utilización de múltiples perspectivas ayuda a construir un cuadro más amplio y completo de la realidad y llama a esto perspectivismo, diferenciándolo del relativismo.

> El perspectivismo asume la necesidad de múltiples contrastes de la realidad como perspectivas para construir un cuadro verdadero, no como un grupo de cuadros verdaderos de diferentes e inconsistentes realidades (Ibíd.: 255, en Shadish, Cook & Leviton, 1991:76).

En tanto que celebra el desarrollo de la lógica y del método científico que garantiza respuestas válidas, es pesimista en cuanto a que éstas puedan ser defendidas para conocimientos tan complejos como es el objeto de estudio de las ciencias humanas: el hombre y los grupos humanos. Por

ello, desarrolla una *lógica probativa* para aplicarla a la lógica evaluativa. Para él, el término *probativo* significa un modelo de razonamiento sencillo y sistemático como el que pueda utilizarse en la elaboración de un juicio sobre algo. Probativo significa contribuir hacia una prueba que no ha sido probada de un modo lógico por su complejidad. En este sentido se puede considerar una prueba para-lógica, que sin ser propio del razonamiento deductivo, probabilístico, las personas pueden creer. Tiene que ver con la práctica, y la ciencia de la valoración debe atender más a la lógica probativa del razonamiento práctico para establecer la plausibilidad de sus hallazgos sobre el valor, que adherirse a un método particular.

Cuestiona a los positivistas por convertir la ciencia del siglo xx, incluida la evaluación, en algo sin valores. Scriven concede una importancia central al hecho de asignar significados de valor a los resultados en los estudios evaluativos.

La evaluación, de acuerdo con Scriven (1983:257), es multicampo en relación con los programas, productos, propuestas, personal, planes y potencialidades; es multidisciplinar, con criterios multidimensionales del mérito; necesita múltiples perspectivas antes de que la síntesis final sea hecha; es multinivel en un amplio rango de niveles de validez, costes, fiabilidad, entre los cuales se debe elegir en orden a ajustarse a los recursos de tiempo y presupuesto; utiliza diferentes niveles de análisis para apoyar la evidencia y documentación apropiada en las diferentes circunstancias, con múltiples metodologías, múltiples presunciones, múltiples impactos y múltiples formatos de informe. La evaluación es la multiplicidad de lo múltiple. Este énfasis sobre la multiplicidad está motivado epistemológicamente para construir un cuadro de la verdad desde múltiples puntos, en coherencia con la epistemología perspectivista de Scriven.

3.2. Las necesidades sociales y la ciencia de la valoración

Scriven considera la evaluación como la ciencia de la valoración y considera que la sociedad demanda la existencia de esta ciencia para utilizarla como un procedimiento sistemático y libre de prejuicios para evaluar los productos y servicios, el personal y los programas e informar si son

buenos, adecuados para las personas que se van a beneficiar de dichos productos y servicios, y lo llama «Evaluación orientada al consumidor».

Ve la evaluación como un servicio para los intereses de todos aquellos que están afectados por la entidad que va a ser evaluada. La evaluación debe cumplir una función de servicio al consumidor. El juicio sobre el objeto evaluado es una de las funciones inexcusables a realizar por el evaluador. Una vez realizada la evaluación, el evaluador tiene una idea clara sobre lo que ha evaluado: ¿Sirve para lo que pretendía servir? ¿Merece la pena que se siga utilizando? ¿En qué debería cambiar para que de verdad sea útil para resolver los problemas o conseguir los objetivo que pretendía conseguir?, se está aplicando bien el programa? ¿A qué se debe que no consiga los resultados pretendidos, al programa? ¿A quiénes lo aplican, por qué? ¿A quienes lo reciben, por qué? Y todo esto va aportando información valorativa, juicios de valor sobre los distintos elementos del programa, de su contexto, de su aplicación, de sus resultados, pero, esos juicios deben ser «objetivos», fundados en la realidad, probados. El juicio del evaluador puede ser: este programa funciona muy bien y merece la pena mantenerlo y extenderlo a otros contextos, usuarios, beneficiarios, pero ese juicio debe estar justificado, documentado, y de ese modo sirve a la formación del responsable del programa, de quienes lo aplican, de los usuarios y de sus beneficiarios.

La evaluación que cumple esta función la denomina «*Evaluación orientada al consumidor*», siendo esencial que todas las partes afectadas por lo que está siendo evaluado, sus necesidades y expectativas sean tenidas en cuenta, e informadas de los resultados de la evaluación (Ibíd.: 249).

3.3. La evaluación libre de costes

La expresión «*evaluación libre de costes*», también atribuida a Daniel Stufflebeam, la justifica Scriven del siguiente modo: «Si bien el interés público demanda evaluaciones, éstas no se justifican en todos lo casos». Y es que las evaluaciones deben atenerse a ciertas normas para que se de un equilibrio de coste-beneficio positivo, de modo que las evaluaciones que crean mas problemas que beneficios deben ser desaconsejadas. El

análisis de la evaluabilidad de lo que se pretende evaluar aconseja cuándo es adecuado realizar una evaluación y cuándo no.

Los evaluadores deben demostrar que los beneficios económicos de una evaluación exceden a los costes de la misma, de lo contrario la evaluación no debe ser hecha (Scriven, 1976: 220). En relación con estas cuestiones, hay que entender que los beneficios no sólo, ni siempre, son exclusivamente económicos, aunque casi siempre tiene relaciones con la economía y el bienestar social, pudiéndose obtener beneficios informativos de claro interés para la mejora de lo que se evalúa, tanto en su ámbito práctico como teórico y político, con el consiguiente beneficio para todo el cuerpo social que se beneficia con ello.

3.4. La evitación de los prejuicios

Scriven quiere minimizar la mayoría de los prejuicios que puedan introducirse cuando se construyen las afirmaciones de valor. No siempre los prejuicios pueden ser evitados por lo que el evaluador debe hacerlos explícitos y establecer alertas contra las dudas o la ignorancia de la existencia de prejuicios. Estas salvaguardias deben ser parte del código que guíe la evaluación en la práctica.

Algunos de los prejuicios estimados por parte de la organización de los centros y de la administración educativa deben tener una protección especial. Ejemplos en relación con los prejuicios son los relacionados con evaluadores que consideran más importante los intereses de los organizadores que otros intereses como los de los propios usuarios o beneficiarios de los programas, o cuando se desarrollan lealtades divididas entre las personas que pertenecen al personal del centro, a las familias, o a la agencia de evaluación. Las soluciones para la evitación de los posibles prejuicios y conflictos de intereses implican la conveniencia de llegar a acuerdos de modo que se garantice la independencia del evaluador.

3.5. La evaluación libre de metas

Un prejuicio frecuente del que debe ser consciente el evaluador es el relacionado con la tendencia a juzgar el programa usando las metas que

los organizadores o el personal responsable del programa mencionan que tiene el programa. El evaluador debe hacer evidentes tanto las metas explícitas como las implícitas de los programas.

El trabajo del evaluador, de acuerdo con Scriven, es identificar cualquier efecto del programa, pretendido o no, que pudiera ayudar a resolver los problemas y mejorar los programas y sus desarrollos aplicativos. Scriven aconseja a los evaluadores que realicen *evaluaciones libres de metas,* también llamadas evaluaciones con referencia a las necesidades a las que el programa pretende dar respuesta o resolver.

> Los evaluadores libres de metas comienzan las evaluaciones totalmente ciegos, como en la justicia, para afirmar las metas de los programas que deben de evaluar. En evaluación estar ciego, como en la aplicación de la justicia, es bello (Scriven, 1973:322).

La evaluación libre de metas hará que el evaluador trabaje con empeño y sin prejuicios por encontrar cualquiera o todos los efectos del programa. Los evaluadores tienen que descubrir los efectos que el programa produce y relacionarlos con las necesidades de aquellos a quienes afecta el programa y para cuya resolución ha sido diseñado. Shadish, Cook y Leviton (1991:81), proponen sustituir los términos evaluación libre de metas por la expresión *evaluación basada en las necesidades y en el descubrimiento de los efectos*. La evaluación libre de metas es similar a la *evaluación respondiente* de Stake (Martínez Mediano, 2007), cuya finalidad es descubrir lo que realmente ocurre en el programa y los beneficios o perjuicios que causa en los implicados y beneficiaros de dicho programa.

3.6. La evaluación formativa y sumativa

Los términos *evaluación formativa y sumativa* están relacionados con dos modos de sensibilidad, o respuesta, a las necesidades de los programas educativos y sociales.

> La evaluación puede ser hecha para proporcionar información a las personas que intentan mejorar algo (evaluación formativa), o para proporcionar información para los que toman decisiones, preocupados por crear, terminar o cambiar algo (evaluación sumativa) (Scriven, 1980:6-7).

La evaluación formativa es fundamental para la mejora continua de lo que se está evaluando (programa, organización, procesos de enseñanza-aprendizaje, etc.), y debe hacerse también, al servicio de la evaluación sumativa, por el mayor alcance de ésta al contribuir a la mejora del programa en general y por los beneficios que puede aportar para la evaluación de programas similares.

Ambas evaluaciones ocupan un lugar importante en las concepciones evaluativas de Scriven, pero él, si tiene que elegir, prefiere la sumativa, que debe ser holística, global. Incluso cuando la evaluación formativa deba ser realizada, debe hacerse al servicio de la evaluación sumativa, de la mejora global del programa para conseguir los resultados pretendidos. «Una buena evaluación formativa, dice Scriven (1974b:9), normalmente implica que hace la mejor simulación posible de una evaluación sumativa». Como vemos, tanto la evaluación formativa como la sumativa tienen funciones de mejora, durante la aplicación del programa o una vez concluido éste (Tabla 1).

Tabla 1. Comparación entre la evaluación formativa y sumativa.

LA EVALUACIÓN SUMATIVA Y FORMATIVA EN LA EVALUACIÓN DE PROGRAMAS		
Dimensiones	**Evaluación formativa**	**Evaluación sumativa**
Propósito	Para mejorar la calidad del programa durante su aplicación. Proporciona información que puede ser compartida rápidamente por aquéllos que están aplicando el programa. Genera informes periódicos. Informar para la evaluación sumativa del programa.	Para probar la calidad del programa. Proporciona información que puede ser utilizada para demostrar a los patrocinados y a la comunidad los resultados del programa. Certificar la calidad del programa. Informar para la mejora del programa y de la organización y sus miembros, para el siguiente ciclo de aplicación del programa.
Audiencia	Administradores y personal del programa.	Usuarios actuales y potenciales del programa.

Dimensiones	Evaluación formativa	Evaluación sumativa
Quién la hace	Evaluador interno.	Evaluador interno y externo.
Cuestiones que hace	Se centra principalmente en las actividades del programa, los resultados, las metas a corto plazo, con propósitos de supervisar el progreso y hacer correcciones durante el curso del programa, cuando se necesite. ¿Cómo funciona? ¿Qué necesita para mejorar?	Se centra principalmente en los resultados inmediatos del programa y en su impacto. Si bien los datos pueden ser recogidos a lo largo de la vida del programa, su propósito es determinar el valor y el mérito del programa centrándose en los resultados. ¿Qué resultados produce? ¿A qué coste? ¿Qué objeciones deben hacerse? ¿Qué cambios se necesitan hacer para mejorarlo?
Exigencias del diseño	Medidas válidas y fiables. Muestreo suficiente. Recogida de información frecuente. Resultados a tiempo.	Medidas válidas y fiables. Muestreo suficiente. Recogida de información al final. Resultados a tiempo.
Aportaciones	Ayuda a aportar sugerencias para la mejora del programa y su aplicación, y para informar al personal que lo aplica.	Ayuda a describir la calidad y la eficacia del programa mediante la documentación de su impacto sobre los participantes y la comunidad.

3.7. La meta-evaluación

La meta-evaluación, término acuñado por Scriven (1967), puede definirse, por analogía con el término evaluación, como una investigación sistemática encaminada a emitir un juicio valorativo acerca de la calidad y los méritos de una evaluación. Scriven enfatiza que los evaluadores tienen la obligación de asegurar que sus propuestas evaluativas o sus evaluaciones están sujetas a evaluaciones competentes. Señala que la meta-

evaluación puede ser formativa, asistiendo al evaluador al diseñar y realizar una evaluación fundamentada, o sumativa, aportando evidencias independientes sobre la competencia del evaluador y de sus informes de evaluación fundados.

Para realizar una meta-evaluación, Scriven señala que el meta-evaluador puede servirse tanto de listas de estándares específicamente elaboradas para la realización de una meta-evaluación concreta, como de los estándares consensuados por la comunidad científica. Él es el máximo promotor de las listas de control de la evaluación y de la meta-evaluación. La meta-evaluación debe aplicarse a la evaluación una vez que se ha realizado, si bien estas mismas normas pueden y deben orientar la planificación y la realización de la evaluación, sirviendo como meta-evaluación formativa e incluso ayudar al diseño, planificación y realización de los programas; si conocemos los criterios que se van a aplicar para valorar un programa y sus resultados, deben guiar su concepción, su diseño y su realización.

La aportación más importante de Scriven en relación con la meta-evaluación es la defensa de su utilización por los evaluadores profesionales con el fin que las evaluaciones y procesos evaluativos sean evaluados y así asegurar la calidad, honradez y precisión del trabajo de evaluación realizado. La meta-evaluación ha sido una de sus muchas preocupaciones para garantizar una evaluación bien hecha y considera que debe ser obligatoria. La meta-evaluación para Scriven es una especialidad evaluativa de singular importancia para la disciplina de la evaluación. Aplicarla es de vital importancia porque representa el cuerpo de conocimiento evaluativo sobre el cual se incorporan las normas para toda la actividad evaluativa, siendo un indicador de que los evaluadores practicamos lo que predicamos, es decir, los beneficios de establecer y defender conclusiones evaluativas con el fin de beneficiar a los ciudadanos, gerentes, prácticos y teóricos (Scriven, 2015, p. 1). La meta-evaluación puede ser formativa, si ayuda al evaluador a planificar la evaluación, y sumativa si le ayuda a realizar una evaluación de la evaluación una vez realizada. Sus informes fomentan la transparencia y orientan el aprendizaje de la comunidad científica.

3.8. La lógica de la evaluación y la práctica evaluativa

Scriven quiere dar consejos prácticos para clarificar la lógica que subyace en la evaluación de modo que pueda mejorarse la práctica. Defiende el razonamiento de la lógica probativa aplicada a la lógica de la evaluación para determinar los resultados. Sobre la evaluación dice:

> La evaluación, comúnmente, implica determinar los criterios del mérito de lo que se evalúa, normalmente desde las necesidades de mejora, y las normas de mérito, frecuentemente como resultado de buscar una comparación apropiada entre lo que se pretende conseguir y lo que consigue el programa tras su aplicación, y después, comparará la práctica del objeto evaluado con las normas dadas (1980:18).

Esta definición de evaluación sugiere tres actividades claras:

1. Determinación de los criterios de mérito o calidad de los programas.
2. Fijación de normas claras sobre lo que debe conseguir la práctica para considerar conseguidos los criterios de mérito de los programas.
3. Medida de la realización, para confirmar que los programas se aplican siguiendo los criterios de su mérito, y el grado con que éstos logran lo que afirman pretender conseguir, de acuerdo con las normas fijadas.

En la *determinación de los criterios del mérito del programa* se deben identificar las dimensiones sobre las cuales el objeto evaluado debe actuar, qué son *las metas y los objetivos* que pretende conseguir el programa.

La *determinación de las normas sobre el mérito* supone señalar cómo deben realizarse esas metas y objetivos para poder considerar que el programa ha trabajado bien para conseguir aquello que el programa pretendía desarrollar. La comparación, en términos absolutos, en relación con los criterios del programa y lo conseguido por sus beneficiarios, es su mérito, y es una evaluación con referencia criterial en relación con aquello que debe desarrollar el programa. Las normas «absolutas» en relación con el nivel de logro en el criterio se deben graduar para concretar el *nivel de logro* en relación con el criterio es *el valor.* De ese modo, podremos decir que los resultados conseguidos por el programa son insufi-

cientes, aceptables, buenos o excelentes, por poner un ejemplo de graduación en el criterio.

La *medida o evaluación de la realización* requiere medir los resultados conseguidos por el programa y compararlos con las normas de mérito. La comparación entre lo deseable, lo posible y lo conseguido nos permitirá realizar *la valoración, o juicio evaluativo*, basándonos en el proceso, el contexto y los resultados.

Finalmente se debe informar de los resultados y de los *juicios evaluativ*os realizados por el evaluador mediante la realización del informe de evaluación hecho de tal manera que sea comprendido por todos aquellos que deban recibir la información, ya sea la dirección del centro educativo, los profesores, los alumnos, la asociación de padres y la sociedad.

4. LA EVALUACIÓN ORIENTADA AL CONSUMIDOR

El enfoque de evaluación orientada al consumidor es, predominantemente, una evaluación sumativa. No obstante, los diseñadores de los productos educativos se han dado cuenta de que utilizar los criterios que defienden los consumidores y las listas de control que los representan cuando se está creando el producto educativo es la mejor manera de prepararse que la subsiguiente valoración del público sea positiva. De ese modo, los criterios y sus listas de control propuestos por las agencias de defensa del consumidor, o de las agencias de la calidad y evaluación de la educación, se han convertido en instrumentos para la evaluación formativa de los productos que se están desarrollando, desde su diseño, antes de su puesta en servicio. Por productos educativos se entiende cualquier elemento de la educación susceptible de evaluar como los nuevos currículos, las nuevas tecnologías, los programas informáticos, el material educativo o cualquier otro servicio educativo.

La evaluación la concibe Scriven como la «determinación sistemática y objetiva del valor y del mérito del objeto evaluado», destacando que la finalidad de la evaluación es siempre la misma: «juzgar el valor» de forma objetiva de lo que se evalúa.

Este enfoque se desarrolló en los Estados Unidos en la segunda mitad de la década de los 60, cuando los nuevos materiales curriculares y otros productos educativos comenzaron a llenar el mercado. Scriven hizo una gran aportación con su distinción entre evaluación formativa y evaluación sumativa:

> La función sumativa de la evaluación permite a los administradores decidir si el currículo totalmente acabado y mejorado por el uso de la evaluación de procesos, en su rol formativo, representa una ventaja suficientemente significativa sobre las alternativas disponibles para justificar el gasto de su adopción por un sistema escolar (Scriven, 1967:41-42).

Los criterios que Scriven sugería para evaluar cualquier producto educativo quedan recogidos en su *Lista de claves de control para la evaluación* (LCE). Las LCE desarrolladas durante años de uso del enfoque centrado en el consumidor para las evaluaciones de los productos, son instrumentos útiles y sencillos para la evaluación, y aportan criterios discriminativos para la selección de los productos educativos disponibles en el mercado.

Las LCE son el resultado de las revisiones encargadas por el gobierno federal de los EEUU centradas en el desarrollo de productos educativos, promocionadas desde los centros de investigación y desarrollo, revisadas y actualizadas continuamente. Los Departamentos de Educación y las Agencias de Evaluación americanos han discutido la utilización de este enfoque de evaluación centrada en el consumidor, dado que más que evaluar los productos y luego comunicar sus hallazgos, o simplemente enumerar los resultados, se han centrado en el uso de formas estandarizadas para recoger y comunicar la información evaluativa sobre cualquier nuevo producto educativo. No obstante, la aplicación del modelo de Scriven plantea un proceso que involucra la valoración de necesidades tanto de los educandos en su contexto próximo, como de la sociedad en la que se desarrolla, mediante la revisión de las metas planteadas y su comparación con otras alternativas, la comparación entre los costes de inversión y su efectividad, la valoración del programa en sí mismo y de los beneficios en personas implicadas. Las guías desarrolladas para este propósito se centran en los contenidos, los procesos educativos, la eficacia y la transferencia.

La información sobre los contenidos se centra en constatar y evaluar la existencia de elementos básicos del programa y de sus resultados, como los siguientes:

— Objetivos claramente formulados.

— Directrices suficientes.

— Otro material necesario.

— Prerrequisitos cognoscitivos, habilidades requeridas.

La información sobre los procesos educativos se centra en comprobar cuestiones como las siguientes:

— ¿Cuál es la naturaleza y frecuencia de las interacciones entre los estudiantes / profesores / administradores / otras personas relevantes?

— ¿Se describe con antelación la estrategia de enseñanza que se va a utilizar de modo que pueda determinarse su adecuación? ¿Ha sido evaluada dicha estrategia?

— ¿Está integrada la evaluación del programa como parte del desarrollo y realización del programa?

La información sobre la efectividad se interesa por las siguientes cuestiones:

— ¿Consigue información rigurosa de los efectos sobre los beneficiarios del programa utilizando uno o más de los siguientes procedimientos: comparación con una base establecida -pretest, postest-; predicción haciendo estimaciones de éxito utilizando técnicas de regresión; comparando con normas locales o del estado; comparación frente a objetivos o normas predeterminadas; comparación de los resultados conseguidos en los estudiantes y/o profesores con la de otros programas o productos alternativos?

— ¿Hay evidencia de que el programa y sus resultados haya eliminado una necesidad documentada tal como ahorro del gasto, mejoras sociales, mayores ratios de aprendizaje?

En relación con la información sobre la transferencia de los resultados se mencionan tres elementos críticos relacionados con la validez ex-

terna, referidos a la transferencia a otros contextos, personas y tratamientos, y utiliza las siguientes cuestiones:

— ¿De qué información se dispone en relación con el aula, el centro y la comunidad en la que el programa o producto ha sido utilizado eficazmente? (tamaño, organización, factores políticos, elementos legales, facilidades, recursos, índices ocupacionales, índices geográficos, factores culturales, factores públicos, factores privados, temas filosóficos).

— ¿De qué información se dispone en relación con los profesores, alumnos, administradores y otras personas relevantes con quienes el programa o producto ha sido utilizado eficazmente?

— ¿Cuál es la edad, grado, experiencia, conocimientos iniciales, expectativas, preferencias, intereses, etnia, cultura?

— ¿De qué información se dispone sobre el tiempo, años en que los programas o productos han sido utilizados eficazmente?

— ¿Tiene el programa, o producto, especiales exigencias en áreas como la formación del profesorado, la organización docente, los materiales y equipos adicionales o la investigación especializada?

Entre las ventajas del enfoque de evaluación orientada al consumidor, podemos decir que este enfoque ha proporcionado un servicio valioso en dos modos:

1. Ha hecho posible las evaluaciones de los productos educativos prestando un servicio a los educadores que pueden no tener tiempo para realizar este trabajo, y
2. ha aportado a los educadores conocimientos sobre los criterios más adecuados para la selección de los productos educativos.

El tiempo y dinero invertido en la comprobación de un producto para el control de la calidad del producto, como es normal, repercute en el precio final, dado que supone invertir tiempo y dinero para valorar y juzgar los programas y productos educativos mediante la realización de estudios piloto antes de que sean puestos a disposición del consumidor. La valoración de los productos educativos en colaboración con las empresas de servicios educativos beneficia a todos.

5. LA LISTA DE CLAVES DE CONTROL PARA LA EVALUACIÓN Y LA META-EVALUACIÓN

La lista de las claves de control para la evaluación de programas, de políticas y para la evaluación de sus evaluaciones, es una herramienta para el evaluador profesional, o práctico, a quien se le supone conocimientos básicos sobre la disciplina de la evaluación y de su vocabulario específico. Esta lista de control se proporciona como un servicio libre al usuario (Scriven, 2013). Su aplicación no tiene por que ser lineal, pudiendo utilizarse con distintos propósitos, como por ejemplo, para el diseño de un programa, para guiar una evaluación o para evaluar la evaluación (meta-evaluación). La evaluación pretende determinar el mérito, valor y significación (m/v/s) del objeto evaluado.

Parte A: Preliminares y metodología

Antes de iniciar una evaluación, conviene responder a cuestiones tales como las siguientes:

— ¿Naturaleza de la petición o necesidad que lleva a realizar una evaluación para conocer el m/v/s o importancia de un programa?

— ¿Funciones de la evaluación: formativa, sumativa, atributiva, o ambas? ¿Debe rendir cuentas, establecer *rankings* o grados, o un perfil de lo evaluado?

— ¿Se piden recomendaciones, explicaciones, críticas, o predicciones a partir de los resultados de la evaluación? ¿Se necesitará ayudar para la utilización, aplicación de la información resultante de la evaluación? Aunque no se pida, debe ofrecerse.

Y, además, hay que identificar:

1. El cliente o persona que oficialmente pide y se compromete a cumplir con los compromisos contractuales derivados de la evaluación, siendo la persona a quien se le entregará el informe.
2. Las posibles audiencias o grupos interesados en conocer el informe.

3. Los promotores.
4. Quiénes, además de los mencionados, tienen derecho a conocer los resultados y la información original producida por la evaluación.

Y concluir con una síntesis informativa.

Es importante acordar antes de iniciar la evaluación, si es factible hacerla y clarificar los compromisos acerca de lo que se va a evaluar mediante una revisión rápida utilizando la LCE. Hay que asegurar la posibilidad de realizar cualquier negociación posterior como consecuencia de los cambios que haya habido que realizar. Y mencionar los reconocimientos y los agradecimientos.

Hay que clarificar los objetivos que orientarán la elección de la metodología de la evaluación a través de cuestiones como las siguientes:

— ¿Se van a utilizar grupos de control o de comparación para determinar supuestos efectos de causalidad del programa en los resultados?
— Si se cuenta con un grupo de control, ¿se pueden asignar los sujetos a los grupos al azar?
— Si se puede seleccionar una muestra, ¿es aleatoria, estratificada...?
— Si no se da ninguno de estos supuestos, ¿cómo se determinarán los efectos de causalidad del programa evaluado?
— ¿Deberá estar la evaluación basada en metas o es una evaluación libre de metas?
— Si hay jueces involucrados, ¿qué control de los prejuicios se utilizará para asegurar la credibilidad y la validez de los resultados?
— ¿Cómo se indagará para saber si hay efectos colaterales diferentes a los pretendidos por el programa?

Hay que identificar los posibles procedimientos de investigación y los análisis de datos que se utilizarán en la evaluación y su justificación. También se debe proporcionar «la lógica de la evaluación», es decir, la justificación general del plan de evaluación.

Parte B: Fundamentos. Aplicación de la LCE

1) Los *antecedentes y el contexto del objeto a evaluar y de su evaluación* que se centran en:

— Identificar la historia reciente y cualquier situación que incida en el programa.

— Identificar a los promotores, personas, grupos y organizaciones que ayudan en la aplicación del programa con fondos u otros recursos.

— Identificar a los clientes o personas que se benefician del programa.

— Identificar cualquier resistencia a la realización de la evaluación.

— Conocer la legislación vigente, las políticas y las actitudes de cambio.

— Identificar la racionalidad subyacente al programa, la teoría oficial del programa y la lógica política, si existe, y si no debe inferirse de un modo fiable.

— Identificar los resultados esperables a la luz de las revisiones de intervenciones similares.

— Identificar las evaluaciones anteriores, si las hubiera, y su impacto, si hubiera.

2) *La descripción de lo que se va a evaluar.* Es importante que el evaluador disponga de la descripción oficial del programa, de sus componentes y de su contexto, y además proporcione una descripción fidedigna y completa del programa real que se aplica para conocerlo y contrastarlo con el oficial, junto con una descripción detallada de las metas, si no es una evaluación libre de metas; debe explicar el significado de cualquier término técnico que no sea del uso de las audiencias y reflejar los modelos significativos e implícitos, las analogías, las metáforas de uso en los participantes; estas descripciones pueden ser más iluminativas que literales, si pueden justificarse; y distinguir entre el programa de intenciones y el de hecho, debiendo evaluar el programa real que se esté aplicando.

3) *Los clientes del programa.* El evaluador debe identificar a los clientes incluyendo a los destinatarios o usuarios de los servicios o productos directos y a los beneficiarios indirectos, por ejemplo, las familias de los

alumnos de un centro educativo, el personal del programa, los que aportan fondos y los responsables de la administración.

4) *Los recursos*. El evaluador debe identificar los recursos financieros, físicos y sociales del programa, incluidos las habilidades, conocimientos, personal, miembros de la comunidad y otros apoyos. E incluir lo que se podría haber usado, no solo lo que se utilizó, definido como el «espacio de posibilidad», lo que podría haber sido hecho como elemento importante para las comparaciones en una evaluación. Puede ser útil listar los recursos específicos que no se usaron o no estaban disponibles para la aplicación del programa, como por ejemplo: ¿hasta qué punto los clientes potenciales, los promotores, los patrocinadores, los voluntarios, no se involucraron como debían haberlo hecho?

5) *Los valores:* Funciones de la evaluación y su finalidad. El evaluador debe identificar los valores relevantes para evaluar el programa de la lista que se aporta más abajo, y comentar la importancia relativa o absoluta de las dimensiones de mérito u otros valores que determine se usarán sobre los hechos del programa para las conclusiones de la evaluación. Estas dimensiones podrían venir valoradas, cualitativamente, en forma de grados, de 1 a 5, o, cuantitativamente, en forma de puntuación del 1 al 10. Además se deben determinar unas normas o estándares mínimos para la aceptabilidad de cada una de las dimensiones de mérito o valor.

El evaluador debe verificar los valores siguientes para determinar su relevancia en el programa, y buscar otros que considere importantes:

1. La satisfacción de las necesidades de la población a la que se dirige el programa mediante comparación con la valoración de necesidades previas a atender. Distinguir las necesidades de actuación de las necesidades de tratamiento, las necesidades previas de las necesidades nuevas encontradas y las necesidades ideales pero poco fáciles de conseguir con los recursos actuales.
2. El criterio de mérito, desde la definición del programa y de su aplicación desde la norma o estándar fijado; pueden incluir puntuaciones del 1 al 10, con un punto de corte para considerar que supera lo mínimo aceptable, en el 5.
3. La legalidad.

4. Requisitos éticos, incluyendo seguridad y privacidad para los implicados.
5. Fidelidad a los supuestos del programa, autenticidad de la aplicación del programa, expresado como índice de aplicación.
6. Metas y deseos personales y organizativos, si no es una evaluación libre de metas.
7. Normas profesionales de calidad que aplica el programa.
8. Requisitos lógicos, por ejemplo, consistencia.
9. La viabilidad científica.
10. La viabilidad tecnológica.
11. El mercado potencial, si fuera relevante.
12. El juicio de expertos.
13. Las normas históricas, tradicionales, culturales.
14. La viabilidad política.
15. La consistencia con el modelo del programa supuesto.

La identificación y aplicación de alguno de éstos criterios es en algunos casos poco relevante, pero en otros será crucial y requerirá consejo del especialista, de los implicados y de los patrocinadores.

Parte C: Sub-evaluaciones

Cada una de estas partes supone realizar una primera fase para confirmar lo que acontece en el programa, seguida de la comparación de los hechos con los criterios de valor, con cómo deberían haber sido conseguidos los hechos.

6) *Los procesos*. Se trata de valorar el m/v/s de todo lo importante que tiene que ser realizado por el programa antes de que los resultados surjan, en relación con las metas pretendidas y que ha podido ser modificado como el plan, la aplicación del programa fidedignamente, la exactitud del nombre oficial, subtítulo o descripción del programa, la dirección, las actividades, los procedimientos, los aprendizajes, las actitudes y los

valores, la moral y la ética, la calidad de la lógica original del programa y su lógica actual en la práctica.

La evaluación del proceso puede incluir la evaluación del rendimiento, de los resultados intermedios que dirigen los resultados finales pretendidos, así como el conocimiento y las actitudes de cambio arraigadas en el personal, profesores, alumnos y otros.

7) *Los resultados.* La evaluación de los efectos positivos y negativos en los beneficiarios directos e indirectos, intencionales y no intencionales, inmediatos, a corto y a largo plazo e imprevistos. La búsqueda de los resultados no siempre puede ser hecha mediante la prueba de hipótesis, porque a menudo los efectos más importantes son los que no se habían previsto. Las metas a corto plazo suelen ser llamadas resultados, o rendimientos. Los resultados pretendidos y obtenidos a largo plazo se conocen como el impacto.

A veces es útil y factible proporcionar explicaciones de éxito y fracaso en torno a los componentes, los contextos y las decisiones en relación con el programa. Hacer esto puede requerir la identificación de la auténtica lógica y teoría sobre cómo debe funcionar el programa en contraste con las lógicas y teorías originales y reales, las oficiales y las implícitas. Los resultados podrían deberse a efectos colaterales que inciden en la población elegida, o tratarse de impactos colaterales que afectan a poblaciones no seleccionadas o a ambas. También se podría indagar sobre posibles beneficios sociales, mejoras medioambientales o resultados más convencionales.

8) *Los costes.* Considere cubrir los costes dinerarios y no dinerarios, directos e indirectos, reales e imprevistos. Detalle en la fase de desarrollo inicial los gastos de partida, de mantenimiento, de progreso para el periodo de tiempo que se prevé dure la evaluación; los costes para componentes tales como salas, equipos informáticos, personal, si es relevante y posible. Incluya el uso de gastos que no siempre se entienden como valor como los costes no dinerarios como el espacio, el tiempo, la experiencia, el trabajo en equipo y el capital político y personal, así como el impacto medioambiental, que raramente es totalmente cubierto por el dinero.

9) *Las comparaciones.* Considere otros medios para conseguir los mismos o similares beneficios a partir de los mismos o menores recursos. Esto se conoce con el nombre de «competidores críticos». Podría

merecer la pena buscar alternativas más económicas casi tan eficaces como otras más costosas. También es recomendable comparar lo evaluado con un enfoque, con otro enfoque ampliamente utilizado que se percibe como alternativa.

10) *La generalización*. En líneas generales, es similar a la «validez externa» de Campbell, pero también cubre sostenibilidad y durabilidad. ¿Puede ser usado el programa con similares resultados en otros lugares, con otro personal, con otros destinatarios, en otros ambientes sociales, políticos, físicos, etc.? La generalización a otros tiempos significa durabilidad y es, casi siempre, crucial. Si este punto de control cubre el coste financiero, social, espacial, temporal, medioambiental, político y otros no monetarios como la capacidad y las condiciones para la supervivencia y la sostenibilidad, esto es aun más importante que la durabilidad, como por ejemplo, cuando se evalúan programas de desarrollo internacional o transcultural.

Parte D: Las conclusiones e implicaciones

11) *Síntesis*. Significado global. Combine las sub-evaluaciones de la Parte C dentro de una evaluación global, por ejemplo, para obtener un perfil global de lo evaluado. El punto de vista debería ser el impacto presente y futuro sobre las necesidades de los clientes beneficiarios de la evaluación, sujeto a las limitaciones éticas y legales y a la viabilidad de otros valores relevantes. Deberá aportar también algunas conclusiones referidas a las necesidades de los clientes y de otros patrocinadores, necesidades de información y deseos, expectativas o esperanzas, si es factible, por ejemplo, las metas que se lograron o los valores no conseguidos.

12) *Las recomendaciones y las explicaciones*. Las micro-recomendaciones como las referidas a gestiones menores y elecciones de equipos de uso son, a menudo, posibles y a ningún coste o esfuerzo extra, y suelen ser muy útiles.

Las macro-recomendaciones que tienen que ver con la disposición del programa total, fondos, costes de continuidad, modificaciones, exportaciones, etc., normalmente requieren:

1. Un conocimiento extenso del ámbito de decisiones por aquellos que deberán tomar las decisiones sobre el programa, quienes no

siempre son los clientes de la evaluación y que pueden no estar dispuestos o ser incapaces de realizarlas.

2. Un esfuerzo extra considerable.
3. Un dominio de la lógica y de la teoría del programa correcto y del espacio de decisión de las partes fundamentales del programa, que a menudo no están accesibles para cualquiera, incluyendo a los expertos y al evaluador.

Las macro-recomendaciones requieren la habilidad de predecir los resultados de los cambios recomendados, algo que una teoría del programa, como muchas teorías en ciencias sociales, a menudo no puede hacer con una alta fiabilidad. Sin embargo, es posible y útil hacer recomendaciones de procedimiento para el futuro, por ejemplo, sobre la necesidad de una investigación más extensa, la recogida de datos más amplia, o cambios en los procedimientos de la evaluación.

También se suelen pedir predicciones sobre cómo funcionará el programa con los cambios recomendados, y esto conlleva riesgos. De hecho, las expectativas de los clientes sobre explicaciones y predicciones son desarrollos para la formación sobre lo que se puede hacer frente a lo que cabe esperar como posible. No obstante, la dificultad de responder a estas expectativas no debe ser una excusa para intentar hacerlas. Nunca puede faltar el requisito fundamental del evaluador de realizar conclusiones evaluativas sobre lo evaluado y que oriente con recomendaciones justificadas sobre cómo aplicarlas para la mejora de los programas.

13) *Responsabilidad y justificación*. Si algo se puede determinar y procede hacerlo, debe ser hecho. Asignar reproches o alabanzas requiere un conocimiento extenso del conocimiento de los principales responsables en la toma de decisiones, de sus recursos y responsabilidades y de un análisis ético de sus opciones y las excusas que pueden tener. No son muchos los evaluadores cualificados para hacer este tipo de análisis, y no deben hacerse a la ligera. Sin embargo, es un hecho evidente que, a veces se cometen errores, y tienen importantes consecuencias, por lo que éstos deben ser identificados y señalados.

14) *El informe y apoyos*. Transmitir las conclusiones de una manera apropiada no necesariamente consiste en entregar un informe técnico de

acuerdo con el modelo de los estudios de investigación clásicos, sino que puede requerir presentaciones radicalmente diferentes para las diferentes audiencias; éstos pueden ser orales o escritos, largos o breves, públicos o privados, técnicos o no-técnicos, gráficos o textuales, anecdóticos o básicos. Se debe incluir un post-informe que ayude y responda a las posibles cuestiones sobre la utilización de los resultados explicando su significado a los diferentes grupos, a los proveedores de fondos y otros beneficiarios. Para este proceso se requiere una amplia variedad de habilidades de comunicación, incluida la idiosincrasia cultural. En este punto se debe abarcar también conseguir resultados y hallazgos de conocimiento incidental, relevantes y coherentes con resultados ya registrados y, si son novedosos, registrarlos, publicarlos y difundirlos para su conocimiento.

15) *La meta-evaluación*. Hay que realizar la evaluación de la propia evaluación para identificar sus fortalezas, limitaciones y otros usos. Esto debe ser hecho siempre por el evaluador como un elemento de control de calidad tan amplio como sea posible después de la evaluación, pero no solo, dado que se puede hacer durante de la realización de la evaluación, antes de realizar el informe final y siempre que sea posible es recomendable hacerlo por un evaluador externo a la evaluación, por un meta-evaluador. El criterio fundamental de mérito de las evaluaciones es la validez, junto con la utilidad —normalmente para los clientes, las audiencias y los promotores—, y la credibilidad, para, a partir del informe de la evaluación informar al personal implicado en el programa. Se puede obtener información relevante mediante la meta-evaluación aplicando la LCE a la propia evaluación por el propio evaluador y después aplicando una lista de control especial para la meta-evaluación como Los Estándares o Normas para la Evaluación de Programas educativos del Joint Committee del 2011.

El uso de la LCE contribuye a que el estudio evaluativo sea comprehensivo y comprensible, creíble y ético, que en él se expliciten los estándares de valor utilizados y que se justifiquen, que sea factible llevarlo a la práctica, suficiente, sensible desde el punto de vista político, que atienda y responda a las necesidades de las audiencias a las que será proporcionado el informe de la evaluación oportuno, seguro, válido y rentable en términos de coste-eficacia. En la Tabla 2 se muestra la LCE.

Tabla 2. La lista de las claves de control de la evaluación de Scriven.

LA LISTA DE LAS CLAVES DE CONTROL DE LA EVALUACIÓN

Parte A: Preliminares y metodología

Naturaleza de la evaluación y su necesidad.

Funciones de la evaluación.

Recomendaciones a partir de los resultados de la evaluación.

Parte B: Fundamentos. Aplicación de la LCE

1. Antecedentes y contexto: del objeto evaluado y de la evaluación.
2. Descripción: qué es lo que se va a evaluar.
3. Clientes: quién comisiona la evaluación y quien se beneficia de ella.
4. Recursos: disponibles para uso del objeto evaluado y de los evaluadores.
5. Valores: Funciones de la evaluación y su finalidad.

Parte C: Sub-evaluaciones

6. Procesos: qué limitaciones, coste, beneficios de las operaciones normales del objeto evaluado.
7. Resultados: qué efectos produce el objeto evaluado.
8. Costes: económicos/ psicológicos/ personales. Costes iniciales/ directos e indirectos/ inmediatos o diferidos.
9. Comparaciones: con opciones alternativas.
10. Generalización: a otras personas, lugares, tiempos y versiones.

Parte D: Las conclusiones e implicaciones

11. Síntesis de todo lo anterior: Significatividad.
12. Recomendaciones y explicaciones: estas pueden ser o no exigidas y pueden realizarse o no a partir de la evaluación.
13. Responsabilidad y justificación.
14. Informe y recomendaciones: Requiere una presentación cuidadosamente estudiada, adaptada a las diferentes audiencias.
15. Meta-evaluación: la evaluación debe ser evaluada, previamente a su realización y a su publicación final y diseminación del informe.

Después de que los datos hayan sido recogidos, analizados y sintetizados el evaluador puede introducir sus opiniones en el informe final. Scriven sugiere que tales tendencias deben hacerse explícitas para corregirlas o evitarlas, si fuera necesario, facilitando un primer borrador del informe para que hagan las críticas oportunas antes de que el informe final sea publicado. En el informe, el evaluador puede incluir recomendaciones sobre el uso o el destino de la evaluación y las posibilidades de generalización de ésta, si fuera posible, a otras personas, programas y situaciones.

La revisión del informe final es un caso especial de meta-evaluación, de modo que se puedan controlar los prejuicios, asegurando que el evaluador o evaluadores sean evaluados. Scriven recomienda que las evaluaciones sean evaluadas utilizando las propias normas de la evaluación, sus conceptos y sus métodos. Para ello los evaluadores pueden usar *La lista de claves de control de la evaluación*.

6. TABLA RESUMEN DEL ENFOQUE DE EVALUACIÓN DE PROGRAMAS DE SCRIVEN

Tabla 3. Síntesis del modelo evaluativo de Scriven.

MODELO DE EVALUACIÓN DE PROGRAMAS DE SCRIVEN	
Elementos para el análisis	**Características**
Definición de evaluación	Valoración sistemática del valor y mérito de las cosas, de lo que se evalúa. Determinación del mérito y valor del objeto a evaluar.
Funciones de la evaluación	Formativa y sumativa.
Concepciones evaluativas	Evaluación orientada al consumidor, evaluación libre de costes, evaluación libre de metas, evaluación formativa y sumativa.
Para quién se hace la evaluación	Para todos los que quieran saber acerca de la calidad de un producto, de un programa, de un objeto, del personal, etc.

MODELO DE EVALUACIÓN DE PROGRAMAS DE SCRIVEN	
Elementos para el análisis	**Características**
Características del evaluador	Experto.
Ámbitos y contextos para su aplicación	Contextos sociales.
Fases principales de aplicación	Determinar el programa. Determinar los criterios del mérito del programa. Determinar las normas sobre el mérito. Medir la realización. Resumir los resultados en un juicio evaluativo.
Metodología evaluativa	Perspectivista. Múltiple.
Información a recoger	Toda la esperada y la no esperada.
Métodos, técnicas e instrumentos	Lógica probativa. Lista de claves de control de la evaluación.
Tipo de informe	Descriptivo.
Posible generalización	La que permite la validez externa matizada por la relatividad de los criterios de valor.
Utilidad de los hallazgos	Para valorar la calidad de los objetos.
Criterios para juzgar las evaluaciones	La meta-evaluación. Listas de las claves de control de la evaluación. Objetividad de los criterios de valor.
Contribuciones	Sus concepciones evaluativas.

EJERCICIOS DE AUTOEVALUACIÓN

1. Defina el concepto de evaluación de programas de Scriven.
2. ¿Por qué Scriven considera «la evaluación como la ciencia de la valoración»? Argumente su respuesta.

3. Aplicaciones de la evaluación formativa y sumativa en la evaluación de programas y su relación con la mejora de las organizaciones y del personal.
4. Concepto de meta-evaluación y funciones de su aplicación.
5. Valore y justifique el enfoque de la evaluación orientada al consumidor de Scriven.
6. Funciones que cumple la lista de las claves de control de Scriven para la evaluación de un programa educativo.
7. Mencione las dos aportaciones más importantes en su opinión de Scriven a la evaluación de programas. Argumente las razones de su elección.

REFERENCIAS BIBLIOGRÁFICAS

Martínez Mediano, C. (2007). Evaluación de programas. Modelos y procedimientos. Madrid: UNED.

Scriven, M. (1967). «The methodology of evaluation». En *Perspectives of curriculum evaluation* (AERA Monograph Series on Curriculum Evaluation. N11) Chicago: Rand McNally.Y en R. Tyler, R.

— (1973). «Goal-free evaluation». En E. R. House (Ed.), *School evaluation: The politics and process* (319-328). Berkeley, CA: McCutchan.

— (1974a). «Prose and cons about goal-free evaluation». *Evaluation Comment*. 3, (1-4).

— (1974b). «Evaluation perspectives and procedures». En J. W. Popham (Ed.), *Evaluation in education: Current application* (3-93). Berkeley, CA: McCutchan.

— (1975). «Evaluation bias and its control». *Occasional Paper,* 4, University of Western Michigan.

— (1976). «Payoffs from evaluation». En C. C. Abt (Ed.), *The evaluation of social programs* (217-224). Beverly Hills, CA: Sage.

— (1980). *The logic of evaluation*. Inverness, CA: Edgepress.

— (1983). «Evaluation ideologies». En G. F. Madaus, D. Stufflebeam y M. Scriven (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation*. Boston: Kluwer-Nijhoff.

— (1986). «New frontiers of evaluation». *Evaluation Practice,* 7, (7-44).

— (1991). *Evaluation Thesaurus*. Newbury Park, CA: Sage.

Scriven, M. (2013). *Key evaluation checklists (KEC)*. Claremont Graduate University & The Evaluation Center. Western Michigan University.

— (2015). *The meta-evaluation checklist*. Claremont Graduate University & The Evaluation Center. Western Michigan University. Tomado el 27 de julio del 2017 de https://wmich.edu/evaluation/checklists

Shadish, W. R. Jr., Cook, T. D. y Leviton, L. C. (1991). «Michael S. Scriven: The Science of valuing». (73-118). En W. R. Shadish, T. D. Cook y L. C. Leviton: O.c.

— (1991). *Foundations of program evaluation. Theories of practice*. Newbury Park, CA: Sage.

Worthen, B. R y Sanders, J. R. (1987). *Educational evaluation: Alternative approaches and practical guidelines.* New York, London: Longman.

Tema 3

El modelo CIPP de Daniel Stufflebeam: La evaluación orientada a la toma de decisiones para la mejora de los programas y de las organizaciones

1. Objetivos y resultados de aprendizaje.
2. Introducción.
3. Antecedentes del modelo CIPP. Las recomendaciones del Comité PDK.
4. La teoría evaluativa y los tipos de decisiones que orienta la evaluación en el modelo CIPP.
   - 4.1. El concepto de evaluación de programas del modelo CIPP.
   - 4.2. Los cuatro tipos de evaluación del modelo CIPP.
   - 4.3. Funciones formativa y sumativa de la evaluación de contexto, de entrada, de proceso y de producto.
   - 4.4. Componentes esenciales de valor del modelo CIPP.
   - 4.5. Filosofía y códigos éticos del modelo CIPP.
   - 4.6. El modelo CIPP como un sistema para la mejora de los programas y de las organizaciones.
5. La práctica evaluativa y las decisiones que orientan los tipos de evaluación del modelo CIPP.
   - 5.1. La evaluación del contexto. Decisiones de planificación.
   - 5.2. La evaluación de entrada. Decisiones de estructura.
   - 5.3. La evaluación del proceso. Decisiones de realización.
   - 5.4. La evaluación del producto. Decisiones de reciclaje.
6. La lista de control para la evaluación.

7. Tabla resumen del modelo CIPP de evaluación de programas de Stufflebeam.

Ejercicios de autoevaluación.

Referencias bibliográficas.

1. OBJETIVOS Y RESULTADOS DE APRENDIZAJE

Con este tema pretendemos que el estudiante conozca el enfoque de evaluación de programas orientado a la toma de decisiones para la mejora de los programas y su gestión, representado por el modelo CIPP de Stufflebeam. Se espera que el alumno, al terminar el estudio del tema, consiga los siguientes resultados de aprendizaje:

— Comprender y valorar el concepto de evaluación de programas del modelo de evaluación orientado a la toma de decisiones de cambio y mejora de los programas y de las organizaciones educativas.

— Comprender y valorar los propósitos de las categorías del modelo CIPP.

— Comprender y valorar las metodologías del modelo CIPP para la recogida de información en la evaluación de programas.

— Valorar la importancia de los criterios y componentes de valor en la evaluación de programas.

— Valorar las aportaciones del 'modelo de evaluación orientado a la toma de decisiones' al desarrollo de la teoría y práctica de la evaluación de programas.

— Valorar y aplicar la lista de claves de control para la evaluación de programas.

— Realizar evaluaciones de programas aplicando el modelo CIPP.

2. INTRODUCCIÓN

El enfoque de evaluación de programas orientado a la toma de decisiones para la mejora de los programas y su gestión está representado por el modelo evaluativo de Alkin (1969), denominado UCLA, por las siglas de la *Universidad de Los Ángeles, California*, y CES (*Centrer for Evaluation Studies*), y el modelo CIPP de Stufflebeam, cuyas siglas representan los cuatro tipos de evaluación del modelo: *Contexto, Insumos o Entradas, Proceso* y *Producto*. Ambos autores renuncian a la evaluación centrada exclusivamente en el cumplimiento de los objetivos del programa, destacando los aspectos claves de la organización de los centros y la toma de decisiones sobre los cambios y mejoras de los programas educativos tras su evaluación. El evaluador y el director organizador responsable del programa trabajan juntos; el primero identifica los aspectos claves que el segundo ha de evaluar aportando información suficiente para valorar su adecuación.

El modelo está basado en el aprendizaje sobre la práctica evaluativa y los esfuerzos continuados para identificar los errores y corregirlos para tenerlos en cuenta en la práctica. La historia del desarrollo del modelo forma parte del desarrollo de los modelos de evaluación desde mediados de la década de los 60 del siglo pasado. Como otros enfoques de evaluación, fue creado porque el clásico modelo de evaluación basado en los diseños de experimentos y en la evaluación basada en el logro de los objetivos, medidos mediante pruebas estandarizados, probaron sus limites para la evaluación de los programas en contextos sociales dinámicos.

3. ANTECEDENTES DEL MODELO CIPP. LAS RECOMENDACIONES DEL COMITÉ PDK

El Comité Nacional de Estudios sobre Evaluación *Phi Delta Kappa* (PDK), en 1969, bajo la dirección de Stufflebeam, recibió el encargo de realizar un estudio sobre el estado de la evaluación. Entre los problemas que fueron detectados destacan los siguientes:

1. La ausencia de una teoría pertinente, de conceptos claros de los requisitos informativos, de los procedimientos apropiados para la planifi-

cación de la evaluación, de los mecanismos e instrumentos para organizar, procesar y presentar la información evaluativa y de personal bien preparado para realizar las evaluaciones.

2. Escepticismo acerca de en qué consiste una evaluación solvente, útil y valiosa que oriente a la toma de decisiones de mejora.
3. La existencia de ambigüedad, frente a la necesaria claridad, en todo el proceso evaluativo.

Los miembros del *Comité PDK* especificaron cinco problemas que debían ser solucionados para mejorar la base conceptual de la práctica evaluativa, que enumeramos a continuación:

1. Necesidad de una *definición clara de lo que se va a evaluar*, con orientaciones metodológicas para la realización de evaluaciones completas.
2. *Evaluación orientada a la toma de decisiones* e información y orientaciones precisas para ayudar a los que toman las decisiones.
3. Elaboración de criterios e instrumentos apropiados para reconocer y tratar *perspectivas de valor* sobre los proyectos y los programas, y su aplicación, y para asignar valor, sobre su calidad y su mérito, a la información obtenida.
4. Reconocimiento de la existencia de *diferentes tipos y niveles de audiencias*, para valorar sus necesidades de información, planificar la recogida de información y de datos y las actividades informativas y, de acuerdo con ello, elaborar *distintos informes* atendiendo a los distintos grupos de interés.
5. *Diseño de la investigación adaptado a las necesidades evaluativas*, por lo que el Comité consideraba que las planificaciones de investigaciones evaluativas debían ser juzgadas mediante normas distintas que las de la mera investigación (Stufflebeam & Shinkfield, 1987:182).

El informe realizado por el Comité PDK concluyó en la publicación del libro titulado *Educational evaluation and decision making* (Stufflebeam *et al.*, 1971). Definían la evaluación educativa como el «proceso de diseñar, obtener y proporcionar información útil y valiosa para juzgar decisiones alternativas» y dividían la evaluación en cuatro actividades diferentes que dan nombre a las siglas CIPP, cada una con sus

propios propósitos de toma de decisiones. En el centro de la evaluación orientada a la decisión subyace el interés por maximizar la utilización de los resultados de la evaluación.

El *PDK* destacó la importancia de la utilización de los resultados de la evaluación más allá de la mera elaboración y presentación del informe sobre los resultados, abogando por que el evaluador orientara a llevar a la práctica las mejoras que debían hacerse en las organizaciones y en los programas a partir de los resultados de la evaluación, en colaboración con la dirección del centro y los implicados en tomar decisiones de cambio en las organizaciones educativas, lo que supuso uno de los primeros intentos en la historia de la evaluación de programas para considerar la evaluación desde una perspectiva centrada en la toma de decisiones.

Una lección importante en torno a los criterios para valorar las evaluaciones fue que las evaluaciones pueden realizarse de un modo incorrecto si se basan exclusivamente en criterios técnicos de adecuación, tales como los requeridos por la validez interna y externa que defienden los diseños experimentales (Campbell & Stanley, 1963). La publicación de 1971 del comité PDK desglosa los criterios de utilidad, junto con el de relevancia, importancia, oportunidad —a tiempo—, claridad y credibilidad, los cuales fueron precursores del trabajo del Comité Conjunto de Estándares para la Evaluación de la Educación, liderado por Stufflebeam, que definió los estándares que deben cumplir las evaluaciones, de utilidad, viabilidad, conveniencia, precisión y responsabilidad.

4. LA TEORÍA EVALUATIVA Y LOS TIPOS DE DECISIONES QUE ORIENTA LA EVALUACIÓN EN EL MODELO CIPP

El modelo CIPP para la evaluación es un marco comprehensivo para realizar evaluaciones sistemáticas de programas, proyectos, personal, productos, organizaciones y sistemas educativos. El modelo fue creado a finales de la década de 1960, para ayudar a mejorar la realización de las evaluaciones de proyectos para la mejora de la enseñanza y del aprendizaje en los centros educativos. El modelo ha ido mejorando con su aplicación tanto en Estados Unidos como en América Latina, Australia, Asia, Europa, y en todo el mundo, para la evaluación de universidades, depar-

tamentos universitarios y colegios. La lista de control para la evaluación del Modelo CIPP es una herramienta que utiliza el modelo, y está disponible en www.wmich.edu/evalctr (Sttuflebeam & Shinkfield, 2007:325).

4.1. El concepto de evaluación de programas del modelo CIPP

La evaluación es definida por Stufflebeam, del siguiente modo:

> La evaluación es el proceso de identificar, obtener y proporcionar información útil y descriptiva acerca del valor y mérito de las metas, la planificación, la realización y el impacto de un objeto determinado, con el fin de servir de guía para la toma de decisiones, solucionar los problemas de responsabilidad y promover la comprensión de los fenómenos implicados (Stufflebeam & Shinkfield, 1987:183).

En esta definición se refleja el propósito de la evaluación, el alcance de la misma, los procedimientos y la utilización de los resultados, todo ello realizado de una forma sistemática y rigurosa.

El modelo CIPP se fundamenta en la evaluación como investigación sistemática acerca del valor de un objeto. Operacionalmente, la evaluación es el proceso de delinear, obtener, presentar informes y aplicar la información descriptiva y de juicio de valor sobre el mérito, valor, significatividad y probidad del objeto evaluado, con el fin de guiar la toma de decisiones, apoyar el rendimiento de cuentas, diseminar las prácticas eficaces e incrementar la comprensión del fenómeno estudiado.

El modelo CIPP recomienda utilizar los estándares profesionales de evaluación tal como son definidos por el *Joint Committee on Standards for Educational Evaluation* (JCSEE, 1981, 1994), (Sttuflebeam & Shinkfield, 2007:326). Los estándares profesionales para la evaluación son principios acordados por los especialistas para realizar y utilizar en las evaluaciones, de modo que sean útiles, factibles, adecuadas, exactas y responsables.

4.2. Los cuatro tipos de evaluación del modelo CIPP

La esencia del modelo viene definida por el acrónimo CIPP que significa la evaluación del contexto, de los insumos o entradas, de los proce-

sos y de los productos o resultados. Para ello, sus autores desarrollaron un marco conceptual para ayudar a los organizadores y administradores a enfrentarse a los *cuatro tipos de evaluaciones* que abarca el modelo, junto con sus correspondientes decisiones, que, en síntesis, consisten en lo siguiente:

1. *La evaluación del contexto* evalúa las necesidades, problemas, recursos y oportunidades para ayudar a los que toman las decisiones a definir las metas y las prioridades, y ayudar a utilizar los juicios sobre las metas, las prioridades y los resultados. Sirve a la toma de decisiones de planificación, que, tras el análisis de las necesidades del contexto, deberá determinar la necesidad de elaborar o seleccionar entre los existentes, un programa educativo, con una clara definición de sus propósitos y de sus objetivos, para atender las necesidades detectadas y definidas.

2. *La evaluación de los insumos, o de entrada*, evalúa los enfoques alternativos de intervención, los posibles planes de acción, incluidos los programas, la dotación del personal, los presupuestos para asegurar la factibilidad, viabilidad y rentabilidad —coste-beneficio— potencial de las propuestas para satisfacer las necesidades específicas y conseguir las metas. Está al servicio de las decisiones de organización, estructurales, que nos informa sobre los recursos disponibles y los que serían necesarios, las diferentes estrategias alternativas para mejorar la aplicación del programa y los planes y propuestas que aportan un mayor potencial para diseñar la planificación que oriente cómo debe ser aplicado el programa.

3. *La evaluación de proceso* evalúa la ejecución de los planes para ayudar al personal a llevar a cabo las actividades y para ayudar a utilizar los juicios valorativos sobre la implementación del programa y para interpretar los resultados que va consiguiendo. Está al servicio de las decisiones de implementación que nos informará sobre cómo se está realizando el plan previsto y las revisiones necesarias que habrá que realizar para mejorar la aplicación del programa.

4. *La evaluación de producto* identifica y evalúa los resultados —intencionales o no, a corto, medio y largo plazo— para ayudar al personal a mantener los programas que se están aplicando en su contexto, en-

focados al logro de resultados importantes y, en última instancia, para ayudar al grupo más amplio de usuarios y beneficiarios a valorar el éxito en los esfuerzos dirigidos a atender las necesidades especificadas. Está al servicio de las decisiones de reciclaje, que nos aportará información sobre los resultados obtenidos, informando acerca de cómo han sido reducidas las necesidades en intensidad, adecuación, y calidad, y, de acuerdo con los resultados, orientará sobre qué se debe hacer con el programa después de su aplicación. En las evaluaciones de resultados a largo plazo, los componentes sobre el éxito de la evaluación pueden dividirse en evaluaciones de eficacia, de sostenibilidad y de transferibilidad. Estas sub-partes de la evaluación de producto preguntan sobre si ¿Fueron los beneficiarios correctamente atendidos? ¿Se respondió a sus necesidades especificadas? ¿Se consiguieron ganancias para los beneficiarios? ¿Fueron sostenibles los mecanismos utilizados para lograr las mejoras? ¿Los procesos que produjeron las mejoras demuestran su transferibilidad y adaptabilidad para su uso eficaz en otro lugar?

En general, estos cuatro tipos y partes de la evaluación de programas deben responder a las siguientes preguntas: ¿qué se necesita hacer?, ¿cómo debe ser hecho?, ¿cómo está siendo realizado? y ¿tuvo éxito en términos de efectividad, impacto, sostenibilidad y transferencia? (Stufflebeam, 2002; Sttuflebeam & Shinkfield, 2007:326-327).

Con ello, se llegó a la *estructura básica del modelo de evaluación CIPP*: la *evaluación del Contexto* como ayuda para la concreción de las necesidades y la determinación de las metas en coherencia con ellas; la *evaluación de Entrada* como ayuda para dar forma a las propuestas y hacer posible su aplicación, teniendo en cuenta los recursos necesarios, incluido el propio programa y la capacitación del personal para llevarlo a cabo; la *evaluación del Proceso* como guía de su realización mediante su aplicación y comprobar si se ajusta a lo planeado y si se consiguen resultados coherentes con las metas, y la *evaluación del Producto,* para informar y valorar los resultados conseguidos por el programa y, junto con los demás informes de procesos, elaborar el informe final al servicio de las decisiones de cambio o continuidad del programa, y de rendición de cuentas.

4.3. Funciones formativa y sumativa de la evaluación de contexto, de entrada, de proceso y de producto

Consistente con el enfoque de mejora, el modelo CIPP coloca la prioridad en orientar la planificación y el desarrollo de la implementación del programa. La función formativa en la evaluación de contexto, entrada, proceso y producto se pregunta respectivamente, ¿qué se necesita hacer?, ¿cómo se está haciendo?, ¿se está realizando con éxito? El evaluador facilita informes a lo largo de la aplicación del programa valorando esas cuestiones para que los patrocinadores estén informados acerca de los hallazgos para ayudar a fundamentar la toma de decisiones y ayudar al personal que aplica el programa en su trabajo hacia la realización y logros exitosos de los resultados. En este orden de cosas el interés del modelo se centra en apoyar a los usuarios de la información de la evaluación, tales como los responsables políticos, administradores y personal del proyecto oportunamente, a tiempo, con información válida para su aplicación, identificando las áreas adecuadas para su desarrollo, formulando metas fundamentadas, planificando actividades y proponiendo presupuestos asociados a los cambios necesarios, a la realización de los planes de trabajo, a fortalecer y mejorar los programas y servicios existentes, a decidir periódicamente si conviene expandir y diseminar los esfuerzos, los programas evaluados, las prácticas efectivas y de ese modo contribuir al desarrollo del conocimiento en el área en la que se está trabajando y responder a los requerimientos de rendición de cuentas de aquellos que financian los proyectos y de todos los grupos legítimamente interesados.

Asimismo, el modelo proporciona guías para realizar la evaluación sumativa retrospectiva al servicio de un amplio rango de grupos de interés entre los que se incluyen a los que financian los programas y servicios, a la organización que los desarrolla y aplica, a las personas que reciben los servicios de esa organización, a los grupos políticos y a los especialistas en programas externos relacionados con el programa que está siendo evaluado y a los investigadores. En el informe sumativo el evaluador incluye aportaciones elaboradas en la evaluación formativa del contexto, entrada, proceso y producto, y utiliza esta información para contestar a las siguientes preguntas retrospectivas: ¿eran importantes las necesidades que pretendía solucionar el programa?, ¿fue el diseño del programa y la planificación de su ejecución acorde con lo que se necesitaba?, ¿se rea-

lizó la aplicación del programa guiada por un diseño correcto y con un presupuesto adecuado?, ¿fue el diseño del servicio, o programa, ejecutado competentemente?, ¿fue modificado si se necesitó?, ¿tuvo este esfuerzo éxito y por qué sí o por qué no? Los usuarios potenciales necesitan respuestas a cuestiones sobre los resultados para ayudar a valorar la calidad, el coste, la utilidad y la competitividad del producto y servicio que ellos podrían adquirir y utilizar (Ibíd.:327).

4.4. Componentes esenciales de valor del modelo CIPP

Los elementos básicos del modelo se sintetizan en tres círculos concéntricos. El círculo central indica los criterios esenciales que deben ser definidos y utilizados en una evaluación. En educación los criterios esenciales son los objetivos educativos que pretende conseguir el programa y debe definirse el grado en el que deben ser conseguidos, su valor. El siguiente círculo está dividido en cuatro sectores correspondiendo a cada foco evaluativo asociado a un programa, o cualquier otra empresa, como son sus metas, su planes de acción, sus realizaciones y sus resultados. El círculo exterior indica el tipo de evaluación a la que sirve cada uno de los cuatro enfoques evaluativos: las metas, a la evaluación de contexto, los planes, a la evaluación de entrada, el desarrollo de las acciones, a la evaluación de proceso y los resultados, a la evaluación de producto (Figura 1).

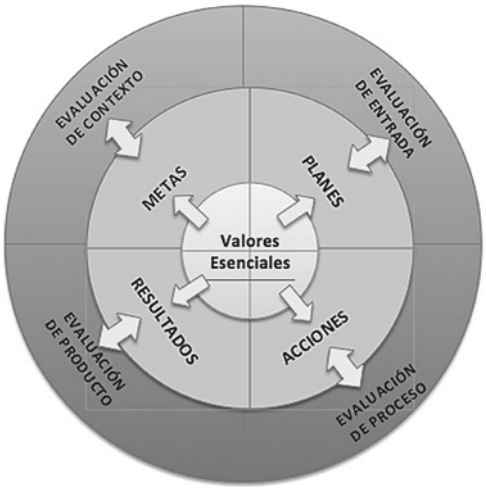

Figura 1. Componentes esenciales de valor del Modelo CIPP.

La afirmación de las metas proporciona cuestiones e indicadores para realizar la evaluación de contexto, para recoger y aportar información para validar y mejorar las metas. La planificación proporciona preguntas e indicadores sobre como conseguir las metas, generando cuestiones para realizar la evaluación de entrada que proporcionará juicios sobre los planes que han sido elaborados para conseguir las metas y cómo mejorarlos. Las acciones que debe realizar el programa para lograr esas metas iniciales aportarán cuestiones para realizar la evaluación de los procesos, las cuales proporcionarán información para la elaboración de juicios sobre las actividades para ayudar a mejorar las estrategias de las personas que aplican los programas. Los logros, la falta de logros y los efectos secundarios atraen la atención de las evaluaciones de producto, dirigen las preguntas sobre la evaluación de producto que en última instancia permiten emitir juicios acerca de los resultados e identificar las necesidades para lograr mejores resultados en las sucesivas aplicaciones de los programas.

La esencia de la evaluación es el criterio y su valor. Los criterios de calidad deben orientar toda evaluación y son, por ejemplo, aquellos que afirman que todos los escolares deben conseguir una formación básica en conocimientos, habilidades, desarrollo personal, desarrollo de la capacidad para continuar el desarrollo personal y profesional a lo largo de la vida, desarrollo del respeto a los derechos humanos, atención a las necesidades específicas de cada niño, al desarrollo de ciudadanos que aseguren la equidad de oportunidades efectivas para las personas, que las familias sean responsables en el desarrollo saludable de sus hijos, que los profesores busquen la excelencia en todos los aspectos, que utilicen los recursos eficientemente, que aseguren la seguridad de los productos y de los procedimientos, que sean respetuosos con las creencias y la religión de cada uno, que apliquen y que realicen investigaciones e innovaciones para mejorar la enseñanza y el aprendizaje, que atiendan responsablemente el derecho a saber de todos. Valores que están en la ciencia y que deben ayudar a describir los criterios y seleccionarlos con ayuda de los implicados y clarificar la información que la evaluación debe atender.

Proporciona las bases para seleccionar y construir los instrumentos de evaluación y los procedimientos y el modo de acceder a la información y definir los estándares de interpretación. El marco de los valores

proporciona un punto de referencia central para detectar las debilidades y fortalezas y orientar las necesidades de mejora de los programas. El examen de un programa frente a un marco de criterios y valores bien construido, podría descubrir que consigue logros positivos inesperados como el apoyo incondicional de la comunidad educativa al proyecto educativo, o la implicación del profesorado en la búsqueda de mejores prácticas, o un apoyo más comprometido de los padres con el proyecto educativo (Ibíd:333-334).

4.5. Filosofía y códigos éticos del modelo CIPP

El modelo CIPP tiene una fuerte orientación de servicio y se basa en los principios de la sociedad libre y los principios democráticos de equidad y justicia. Implica a todos los grupos legítimamente interesados. Se orienta a la mejora de los programas y se centra en realizar evaluaciones objetivistas. La orientación epistemológica del modelo CIPP es objetivista, no relativista. El diccionario de la Real Academia de la Lengua Española define el término objetivismo como 'cada uno de los planteamientos científicos o filosóficos que sostienen la realidad de valores y verdades con independencia del sujeto.'

Se pide a los evaluadores y a sus clientes identificar a los grupos de interés relevantes, a los legítimos beneficiarios de la evaluación, e implicarlos, o al menos a sus representantes para alcanzar consensos que ayuden a afirmar los valores fundamentales, definir las cuestiones de la evaluación, clarificar los criterios evaluativos, contribuir a las necesidades de información y ayudar a clarificar las necesidades del servicio y de información y, finalmente, valorar el mérito, el valor, el significado y la adecuación del servicio y ayudar a interpretar los hallazgos y favorecer su utilización.

La idea central de las evaluaciones de acuerdo con el modelo CIPP es proporcionar información fundamentada que pueda ayudar a los responsables del programa o servicio a recoger información evaluativa regularmente para mejorar el programa, hacer uso efectivo y eficiente de los recursos, a tiempo, con el fin de servir a las necesidades especificadas en el programa e informar a los beneficiarios adecuadamente y con equidad (Ibíd:330).

Un principio fundamental del modelo es que el propósito más importante de la evaluación no es probar sino mejorar. Aquéllos programas que no merezcan la pena ser mejorados deben ser desacreditados y aconsejada su no utilización. Los evaluadores deben identificar el valor de los programas, tanto los que tienen éxito como los que fracasan en sus esfuerzos para aportar soluciones a necesidades y problemas.

Se basa en los principios éticos de la Declaración Universal de los Derechos Humanos de Naciones Unidas. Advierte sobre la importancia de controlar los prejuicios y los conflictos de interés para realizar las evaluaciones y alcanzar conclusiones. Defiende y justifica el establecimiento de normas técnicas de mérito, obtener y validar los hallazgos desde múltiples fuentes, buscar las mejores respuestas, organizar de la mejor manera las conclusiones e informar de los hallazgos con honestidad y justicia atendiendo el derecho a conocer, a saber de las audiencias, actuar en la evaluación de un modo independiente e identificar las necesidades para futuras investigaciones. Las evaluaciones objetivistas tienen como intención alcanzar las conclusiones correctas no aquellas que interesan a un determinado grupo.

El modelo CIPP pide a los evaluadores cumplir con los estándares profesionales de la evaluación de programas. Los principales estándares a los que invoca requiere que las evaluaciones cumplan con los requerimientos de las normas profesionales de utilidad, factibilidad, adecuación y exactitud o precisión y de rendimiento de cuentas. Además, los evaluadores deben utilizar la meta-evaluación formativa para guiar el trabajo de evaluación y corregir deficiencias mientras la realiza, y la meta-evaluación sumativa. En el informe final de la evaluación, los evaluadores deberán explicar sus juicios de valor y justificar el grado en que la evaluación cumple cada uno de los estándares relevantes del Joint Committee (JCSEE). Cuando sea posible, la evaluación debe estar sujeta a una meta-evaluación externa independiente, lo que ayudará a resolver cualquier conflicto de interés reflejado en el procedimiento y en el informe realizado. El informe del meta-evaluador externo debe estar disponible para todos los miembros y audiencias que tengan derecho a saber. Los evaluadores deben indicar y explicar sus juicios sobre el grado en que la evaluación cumplió con cada una de las normas pertinentes y relevantes del JCSEE e informar sobre los logros y lo que no han sido conseguidos junto con los

efectos secundarios, para así atraer la atención sobre la evaluación del producto, dado que para emitir juicios es esencial una valoración sobre los resultados e identificar y señalar qué se necesita hacer para lograr mejores resultados. Todo ello debe ser comunicado a todos los miembros de la organización para conseguir una comprensión compartida sobre los puntos fuertes de la organización y las debilidades, las necesidades, los recursos, las oportunidades y priorizar los problemas que deben ser atendidos (Ibíd:332).

4.6. El modelo CIPP como un sistema para la mejora de los programas y de las organizaciones

El modelo CIPP considera la evaluación como un elemento indispensable para la mejora y rendición de cuentas en un marco de valores apropiado y en la búsqueda de respuestas claras e inequívocas a las metas planteadas. Está configurado especialmente para guiar la realización de análisis exhaustivos y sistemáticos que ayuden a mejorar proyectos y programas dirigidos a resolver problemas tal como se desarrollan en contextos reales, no en las condiciones controladas que estudian los diseños experimentales.

El modelo considera la evaluación esencial para el progreso social y el bienestar de los individuos y grupos. Los grupos sociales no podrán mejorar sus programas, servicios y productos a menos que aprendan por qué fallan y por qué consiguen sus éxitos. Los desarrolladores y proveedores de servicios no pueden asegurar que las metas de los programas son valiosas a menos que esté validada su consistencia con los criterios fundamentados en respuesta a las necesidades definidas de los beneficiarios o grupos a los que se dirige.

Los desarrolladores y los proveedores de programas y servicios no pueden planificar de manera efectiva e invertir su tiempo y recursos de manera inteligente sin previamente identificar y valorar diferentes opciones. No pueden conseguir credibilidad, respeto y apoyo si no demuestran que los proyectos, programas, planes planteados y aceptados han sido realizados con responsabilidad y han producido resultados beneficiosos. No se puede construir sobre las experiencias realizadas si no se

observan, estudian y se actúa a partir de las demostraciones sobre los esfuerzos fallidos y los exitosos. Y no podrán convencer a los consumidores para que adquieran sus servicios y productos a menos que se basen en su valor demostrado e informado transparentemente y con honestidad del servicio.

El personal de una institución no puede valorar que un programa ha dado respuesta satisfactoriamente a todas las necesidades definidas si no desarrolla la capacidad de su personal para llevar a cabo evaluaciones internas y si no contrata a un evaluador para la realización de evaluaciones externas objetivas, y los evaluadores no podrán defender sus conclusiones evaluativas a menos que se basen en información fundamentada y en criterios defendibles. Asimismo, los evaluadores internos y externos no podrán sostener la credibilidad de sus evaluaciones si no son sometidas a meta-evaluaciones basadas en normas apropiadas como las del JCSEE (Ibíd.:351-352).

5. LA PRÁCTICA EVALUATIVA Y LAS DECISIONES QUE ORIENTAN LOS TIPOS DE EVALUACIÓN DEL MODELO CIPP

La evaluación es un proceso en tres etapas para identificar, obtener y proporcionar información para la toma de decisiones. Stufflebeam defiende que son tres los principales propósitos de la evaluación: 1) Promover la comprensión de los fenómenos implicados. 2) Guiar la toma de decisiones. 3) Informar para la responsabilidad.

El objeto a evaluar debe de ser valorado en los siguientes aspectos: 1) sus metas, 2) su planificación, 3) su realización, 4) sus resultados y su impacto.

El criterio de la evaluación debe de tener en cuenta: 1) su valor, en las respuestas a las necesidades detectadas y 2) su mérito, en relación con la calidad con la que son resueltas dichas necesidades.

Para proporcionar información a quienes toman decisiones fueron identificados los principales tipos de decisiones que afrontan para, de acuerdo con ellas, crear las estrategias evaluativas necesarias. De este modo, se identificaron y concretaron los *cuatro tipos de decisiones* que

orienta la práctica educativa. El modelo emplea varios métodos, se basa en una amplia gama de técnicas para la recogida de información y es apoyado por una extensa literatura teórica y práctica.

Conviene decir que la primera vez que se planifica y realiza la evaluación de un programa es la más laboriosa.

5.1. La evaluación del contexto. Decisiones de planificación

Referir la evaluación de contexto exclusivamente a la evaluación de necesidades supone no analizar cuestiones sobre problemas, recursos, aspectos positivos y oportunidades. Estos cuatro elementos son muy importantes para diseñar proyectos, programas y servicios bien fundamentados, no exclusivamente basados en las necesidades, lo que supondría centrarse de un modo muy estrecho en el problema o problemas a resolver.

Los principales objetivos de los estudios de evaluación de contexto son la valoración del estado global del objeto a evaluar, ya sea un proyecto, programa, personal, instituciones y/o sistemas educativos, para identificar sus deficiencias y virtudes, diagnosticar los problemas cuya solución puede mejorar el estado del objeto, caracterizar el marco en el que se desarrolla el objeto o programa a evaluar y si las metas y prioridades existentes que afirma perseguir el programa están en consonancia con las necesidades que debe satisfacer.

Para realizar la evaluación de contexto suelen utilizarse metodologías tipo encuesta para valorar las necesidades de los grupos a los que se pretende atender, mediante cuestionarios específicos y entrevistas para preguntar a los que van a desarrollar y a aplicar los programas y comprobar si las metas fijadas son coherentes con las necesidades afirmadas. El evaluador puede analizar estudios previos para la valoración de necesidades de los grupos de interés y contrastar la información con los promotores de los programas y demás grupos implicados. También se pueden aplicar pruebas de conocimientos o de otro tipo para realizar la evaluación de necesidades, así como utilizar la técnica Delphi para ampliar la visión consultado a especialistas y grupos de discusión para conseguir acuerdos y priorizar las necesidades, todo ello dependiendo de la temática a evaluar, el contexto o los grupos a atender.

Las evaluaciones de contexto pueden ser iniciadas antes, durante o después de la aplicación de un proyecto, programa o cualquier otra intervención. Cuando la evaluación de contexto se realiza durante o después de que el programa haya sido aplicado se aconseja realizarlas en combinación con las evaluaciones de entrada, proceso y producto. En este caso la evaluación de contexto es útil para ayudar a valorar, juzgar, si las metas establecidas facilitan a los grupos interesados si los esfuerzos han merecido la pena para atender a las necesidades de los beneficiarios y otros grupos de interés.

Las evaluaciones de contexto tienen un amplio rango de uso y orienta la toma de decisiones para la selección de nuevas metas. Una evaluación de contexto debe proporcionar información para que un director, administrador o responsable de una institución comunique de un modo fundamentado a los miembros de la organización y demás grupos de interés, la conveniencia o necesidad de abordar un proyecto aportando el estudio de los puntos fuertes de la organización y sus debilidades, de las necesidades y describir sus metas y objetivos, las oportunidades y los problemas prioritarios que podrían solucionar dicho proyecto, innovación, programa o servicio. Y dicha información podría utilizarse para buscar apoyos financieros externos u otro tipo de apoyos y contratos. Cuando los programas y sus acciones se están aplicando, la evaluación de contexto puede servir para asegurar que dichos programas se centran en las necesidades para las que se han creado.

La evaluación de contexto debe analizar también los problemas o dificultades que habrá que superar para responder a las necesidades afirmadas, conocer y valorar los *recursos* de que se dispone, los profesionales y su preparación en relación con el programa, y los *medios y apoyos* a los que se puede recurrir en el ámbito local que pueden ayudar a apoyar los esfuerzos y a conseguir los propósitos. La *concreción de propósitos defendibles*, asumibles, relacionados con la misión de la institución, los propósitos que suscriben y los estándares éticos y legales.

La evaluación de contexto debe proporcionar una base sólida para el ajuste de las metas y prioridades y para la indicación de los cambios necesarios que debe conseguir el programa, y que deben reflejarse en esta primera parte de la evaluación. Su metodología es la propia de la investigación tipo encuesta, y utiliza instrumentos tales como los cuestionarios,

las entrevistas, la observación, las reuniones sobre temas claves para el debate, la técnica Delphi, grupos de discusión, lecturas de informes, etc., para la recogida de información destinadas a generar propuestas acerca de los programas de intervención más adecuados, para generar los cambios conducentes a la solución de las necesidades detectadas a las que debe servir.

Registrar estas necesidades, sus fundamentos, las metas a desarrollar y como conseguirlas es fundamental en las organizaciones para realizar un seguimiento de sus progresos y revisar el estado de la situación periódicamente. Es importante registrar la historia de anteriores evaluaciones de contexto, sus decisiones, desarrollos y resultados y analizarlos para ayudar a la toma de decisiones, evitando procedimientos poco exitosos y valorando los que aportaron mejores resultados (Ibíd.:336-337). La evaluación del contexto se utiliza al servicio de las decisiones de planificación que ayudarán a seleccionar, formalizar y concretar las metas del programa y a revisar su aplicación y sus resultados.

5.2. La evaluación de entrada. Decisiones de estructura

Las *Decisiones de estructura* hacen referencia a los medios necesarios para alcanzar cada una de las metas o necesidades valoradas. Estos medios estarían representados por los propios programas, los recursos necesarios para su aplicación, los presupuestos, junto con la elección del plan concreto y los procedimientos para llevarlos a la acción. Stufflebeam lo llamó decisiones de estructura y deben estar fundamentadas en las evaluaciones de entrada que ayudarán a identificar y valorar los méritos relacionados con las planificaciones de diversos proyectos y programas acordes con las necesidades aportadas en la evaluación de entrada. Ayudan a seleccionar el programa que mejor contribuya a solucionar las metas definidas para responder a las necesidades analizadas en la evaluación de contexto. Se deben analizar los potenciales programas y sus enfoques y requerimientos para su aplicación, incluidos los ya aplicados y evaluados sus resultados, así como la calidad de los mismos, sus necesidades de recursos y tiempos de aplicación y los resultados que pueden conseguir y su impacto en el contexto en el que se pretenden aplicar.

También debe identificar y valorar las estrategias aplicables y explicarlas. Asimismo debe delimitar el marco contextual del estudio y conocer las limitaciones y los recursos potenciales utilizables que deben tenerse en cuenta en el proceso de desarrollo de la aplicación del programa.

La metodología de investigación y recogida de información utilizada debe centrarse en la revisión del estado de la práctica en relación con la satisfacción de las necesidades específicas, incluida la valoración de programas y estrategias que se han diseñado para solucionar problemas similares, la valoración de estrategias adecuadas para conseguir las soluciones, la escucha del personal para que exprese sus necesidades, intereses y valoraciones para así poder hacer una estimación realista de los recursos y dificultades que serán necesarios afrontar durante el proceso de aplicación del programa de intervención para dar respuesta a las necesidades definidas.

Las evaluaciones de entrada suponen diversas etapas. Pueden basarse en el análisis de estado de la práctica sobre cómo se consiguen las metas pretendidas por si fuera necesario revisar los planteamientos previos. Mediante un grupo de discusión se pueden valorar las necesidades, estrategias, información de aplicaciones realizadas, enfoques y procedimientos diferentes en la búsqueda de soluciones, su viabilidad y los resultados que consiguen. Las conclusiones de este grupo de discusión pueden aportar información que ayuden a decidir sobre las decisiones más convenientes, propuestas presupuestadas para realizar peticiones a diversas administraciones o fundaciones para que financien dichas propuestas. También se puede contar con un grupo de trabajo para diseñar las estrategias que debe realizar el programa, implicando al personal que lo va a llevar a la práctica, y orientar sobre cómo informar e interesar a los profesores en la conveniencia de la aplicación del nuevo proyecto y su programa y sobre cómo utilizar los recursos disponibles.

Realizar informes de la evaluación de entrada a los grupos de interés explicando las alternativas posibles y sus potenciales resultados y evitar pérdida de tiempo y recursos con innovaciones que han demostrado resultados inadecuados, es importante para ayudar a la toma de decisiones, contribuir a la transparencia y rendición de cuentas, y conseguir ayudas financieras que contribuyan a realizar los esfuerzos de cambio (Ibíd.: 338-340).

5.3. La evaluación del proceso. Decisiones de realización

La evaluación de proceso debe comparar las actividades que está realizando el programa con las planeadas y los gastos que tiene en relación con lo presupuestado, describir los problemas de implementación que encuentra y valorar cómo el personal lo aplica. Tiene como finalidad la comprobación de la implementación del plan y debe documentarlo para ayudar a las decisiones de realización sobre cómo se están aplicando los proyectos y programas planificados. El evaluador debe hacer un plan lo más respetuoso posible para que no interfiera en las actividades que realiza el programa, debe observar las actividades del personal implicado en el desarrollo del programa, en su aplicación. Debe presentar informes sobre los resultados de la evaluación de proceso cada poco tiempo a los grupos que deben recibir la información, con el fin de que puedan detectarse, señalarse y solucionarse los problemas de realización de los programas a tiempo.

Además de un evaluador profesional es importante designar a una persona de la organización para que haga un seguimiento de los procesos y recoja información y la documente, y realice juicios valorativos sobre la calidad con la que se está aplicando el programa para evitar desviaciones importantes, orientar su mejora y reducir costes. Antes ha debido revisar el plan de trabajo que debe realizarse, las estrategias, las actividades sobre las que debe realizar un seguimiento y debe desarrollar un plan de actividades para la recogida de información sobre el programa que se está implementando.

Algunos de los objetivos de la evaluación del proceso son:

— Comprobar el ritmo de las actividades en la aplicación del programa.

— Observar si el programa se desarrolla tal como se había planeado y a qué puede deberse en aquello que no lo esté haciendo.

— Comprobar si se utilizan los recursos disponibles de un modo eficiente.

— Guiar para explicar o modificar el plan que se había previsto cuando fuera necesario.

— Valorar periódicamente hasta qué punto los que participan en el programa aceptan y son capaces de desempeñar las funciones que le están encomendadas.

— Proporcionar un informe extenso del programa que realmente se está llevando a cabo, compararlo con el que realmente se había planificado, establecer una valoración objetiva de los costes y juzgar la calidad del trabajo, tanto por los evaluadores como por los participantes implicados en el desarrollo del programa.

En cuanto a la metodología, el evaluador del proceso debe revisar el plan del programa, y cualquier evaluación anterior, para identificar aspectos importantes del programa que deben ser controlados. Puede hacerlo a través de las siguientes situaciones:

— Reuniones de trabajo con el personal implicado en el programa.

— Desarrollo y puesta en marcha de materiales.

— Revisión de la planificación del trabajo.

— Observación y seguimiento de la enseñanza en el aula.

— Recomendaciones para los estudiantes.

— Valoración de las habilidades desarrolladas en los estudiantes en relación con el programa.

— Revisión de la planificación y realización de trabajos en casa por los alumnos.

— Utilización de tests diagnósticos.

— Utilización de tests de rendimiento.

El evaluador del proceso debe desarrollar un esquema general de las actividades para la recopilación de datos y llevar a cabo la recogida de información. Puede trazar un boceto o esquema acerca del funcionamiento del programa tras visitar y observar los centros de actividad, revisar los documentos del programa, asistir a las reuniones y entrevistar a los principales representantes en relación con el programa. Luego podrá preparar un informe que resuma el plan de recogida de información y datos y después analizará cómo se está realizando la aplicación y los problemas claves que se detectan.

La evaluación del proceso es una fuente de información vital para interpretar los resultados de la evaluación del producto, o resultados finales del programa.

Una vez elaborado el informe de evaluación de proceso deben realizarse reuniones grupales para exponer los resultados y valorar las estrategias de mejora a tiempo. El informe puede presentarse al personal en una reunión para discutir los resultados de proceso con los implicados en el programa y utilizarlo para una revisión del programa y de los planes del evaluador para seguir recogiendo información y datos, pidiendo sugerencias a todos para mejorar la recogida de información a través de cuestionarios, entrevistas, observaciones, diarios o grupos de discusión, pruebas diagnósticas y pruebas de rendimiento.

Y pueden proyectarse futuras sesiones para el intercambio de información. El evaluador debe mostrar que el propósito principal de la evaluación de un proceso de aplicación de un programa es ayudar al personal a llevar a cabo, de la mejor manera posible, su programa, a través de una especie de proceso de monitorización y cualificación del personal que lo aplica. Estos informes contribuyen a la realización del informe final para el rendimiento de cuentas (Ibíd.:341-344).

5.4. La evaluación del producto. Decisiones de reciclaje

Si el propósito de una evaluación del producto es valorar, interpretar y juzgar los logros de un programa, su objetivo es averiguar hasta que punto el programa ha satisfecho las necesidades del grupo al que pretendía servir en aquello que se había propuesto lograr y qué otras aportaciones ha conseguido a lo largo de la aplicación del programa. Los evaluadores deben trabajar con el personal que aplica los programas en la evaluación y valoración de los resultados.

La evaluación del producto debe ocuparse de los efectos del programa, tanto de los deseados como de los no deseados, de los positivos y de los negativos. Debe realizar mediciones de los resultados conseguidos, interpretarlos y elaborar juicios sobre ellos y aportar un feedback formativo.

> Una evaluación del producto debe, por lo general, examinar los resultados desde muchos puntos ventajosos: en conjunto, por subgrupos o individualizado. Un resultado relacionado con un individuo puede ser clasificado como un éxito o como un fracaso según haya satisfecho o no

> una necesidad previamente diagnosticada del individuo (Stufflebeam & Shinkfield, 1987:201).

Debe valorar los resultados de la aplicación del programa y compararlos con las necesidades previas definidas. La información sobre los logros y el modo de mejorarlos es importante realizarlo durante todo el ciclo de vida del programa y también al final de su aplicación, y se debe extender la evaluación de resultados a largo plazo, a la evaluación de impacto. Debe comparar los resultados conseguidos con los esfuerzos realizados, personales y de recursos económicos, y debe valorar el mérito y el valor de los conseguidos y comparar estos resultados con los que consiguen programas similares. Los resultados deben ser analizados por grupos de cuestiones importantes pretendidos por el programa y también en relación con la totalidad así como los beneficios aportados para los diferentes grupos y para la organización en su totalidad.

Y debe utilizar las metodologías y las técnicas e instrumentos más adecuadas a cada objetivo de la evaluación, o combinación de técnicas a lo largo de todo el proceso de evaluación, lo que ayuda a una mejor comprensión de cómo ha funcionado el programa para conseguir los resultados logrados, medidos mediante escalas de valoración, pruebas de logro o de rendimiento y cuestionarios de satisfacción.

El análisis de los resultados puede generar hipótesis para continuar con investigaciones de casos concretos que requieren de nuevos proyectos y programas. Los evaluadores deben interpretar los resultados conseguidos y explicar si unos determinados resultados se deben al modo como ha sido aplicado el programa, a sus recursos, a la planificación inicial o a cualquier otra causa.

El informe sobre los resultados de una evaluación del producto puede desarrollarse en distintas etapas. Los informes de *final de ciclo* pueden resumir los resultados alcanzados e interpretarlos a la luz de las necesidades preevaluadas, los costes en tiempo, recursos y personal, y los puntos del plan que han sido conseguidos. El informe de los resultados debe proporcionar una guía para modificar el programa mejorándolo en aquello que no ha funcionado bien. La evaluación de programas concluye en un informe para ayudar a los grupos a planificar, realizar, institu-

cionalizar y difundir un servicio efectivo en beneficio de la comunidad educativa y de la sociedad.

A partir del informe completo, los responsables y grupos interesados deben reflexionar sobre sus resultados y valorar si es necesario aportar información complementaria para matizar algunos de los hallazgos. A partir de la valoración sobre el programa, su funcionamiento y los resultados que consigue en respuesta a unas necesidades definidas, se deben tomar las decisiones que corresponda sobre la conveniencia de continuar aplicando el programa, aplicarlo en otros contextos, introducir las mejoras que aconseja el informe o decidir acerca de la aplicación de un nuevo programa.

La evaluación del producto es un componente esencial para el informe de rendición de cuentas. Un informe documentado sirve para convencer de la conveniencia de seguir aplicando el programa, y de la necesidad de seguir financiando su aplicación, o incluso proporcionar financiación y apoyos adicionales, o, por el contrario, considerar que dicho programa no merece la pena que siga siendo financiado (Sttuflebeam & Shinkfield, 2007:345-346).

Las *decisiones de reciclaje* se relacionan con la conveniencia de continuar con la aplicación del programa en el siguiente curso o decidir no volver a aplicarlo, unirlo a otro nuevo programa como refuerzo o institucionalizarlo para un período determinado de tiempo y buscar fondos, si fuera necesario, para el curso siguiente, basándose en los informes sobre la realización del mismo. La evaluación del producto tiene implicaciones psicológicas: refuerza el trabajo bien hecho, tanto en el personal que ha aplicado el programa como en los destinatarios o beneficiarios del mismo, o reduce el entusiasmo cuando los resultados no han sido los esperados.

Cada tipo de evaluación, en el modelo CIPP, desempeña funciones únicas, y puede utilizar diferentes métodos, pero existe una relación simbiótica entre ellas, una apoya a la siguiente y aporta información fundamental para comprender la totalidad del proceso evaluativo (Stufflebeam & Shinkfield, 1987: 203). En la tabla 1 se reflejan los objetivos, métodos y toma de deciones a las que sirve cada uno de los cuatro tipos de evaluación.

Tabla 1. Los cuatro tipos de evaluación (Stufflebeam y Shinkfield, 2007:335)

	Objetivo	**Método**	**Relación con la toma de decisiones en el proceso de cambio**
Evaluación de Contexto	Definir el contexto pertinente, identificar la población objeto y evaluar sus necesidades, identificar oportunidades para abordar las necesidades, diagnosticar los problemas subyacentes a las necesidades y juzgar si los objetivos del programa son suficientemente receptivos a las necesidades evaluadas.	Utilización de análisis de sistemas, encuesta, revisión de documentos, análisis de datos secundarios, dialogo, entrevistas, pruebas diagnosticas y técnica de Delphi.	Para decidir sobre las situaciones a las que sirve; Las metas asociadas a las necesidades o aprovechamiento de las oportunidades; Las prioridades para los presupuestos y recursos a tiempo; Los objetivos asociados con la solución de problemas, para planificar los cambios necesarios; proporcionar una base para juzgar los resultados.
Evaluación de Entrada	Identificar y evaluar las capacidades de los sistemas, estrategias de programas alternativos, diseños de procedimientos para implementar las estrategias, presupuestos y horarios.	Inventariar y analizar recursos humanos y materiales disponibles, estrategias de solución de problemas y diseños de procedimientos viables relevantes, costo beneficio; Utilización de métodos tales como búsqueda bibliográfica, visitas a programas ejemplares, equipos de asesores y ensayos piloto.	Para seleccionar recursos de apoyo, estrategias de solución y diseños de procedimientos, es decir, para estructurar actividades de cambio y presupuesto y programar las actividades del programa y proporcionar una base para juzgar la implementación.

Evaluación de Proceso	Identificar o predecir defectos en el diseño de procedimientos o en su implementación, proporcionar información para las decisiones pre-programadas, y registrar y juzgar los eventos y actividades de procedimiento	Monitorear las posibles barreras de procedimiento de la actividad y permanecer alerta ante las imprevistas, obtener información específica para las decisiones programadas, describir el proceso real e interactuar continuamente y observar las actividades del personal del proyecto y otras partes interesadas.	Para implementar y refinar el diseño y los procedimientos del programa, es decir, para efectuar al control del proceso y proporcionar un registro del proceso real para su uso posterior en la interpretación de los resultados.
Evaluación de Producto	Recopilar descripciones y juicios de los resultados y relacionarlos con los objetivos y con el contexto, entradas y la información de proceso; y para interpretar su mérito, valor, importancia y probidad.	Mediante la definición operacional y la medición de los criterios de resultados, la recolección de juicios sobre los resultados de las partes interesadas, la realización de análisis cualitativos y cuantitativos y la comparación de los resultados con las necesidades evaluadas.	Para decidir continuar, terminar, modificar o reorientar la actividad de cambio; y para presentar un registro claro de los efectos (intencionales y no intencionales, positivos y negativos) en comparación con las necesidades evaluadas y los objetivos especificados.

6. LA LISTA DE CONTROL PARA LA EVALUACIÓN

La lista de control es una herramienta ajustada al Modelo CIPP de evaluación de proyectos, programas, personal, instituciones y sistemas, y tiene como finalidad ayudar a los evaluadores y responsables de los programas en la evaluación de programas y organizaciones con el objetivo de conseguir mejoras sostenibles a largo plazo. No obstante, el modelo CIPP aplica las Normas del *Joint Committee*.

Correspondiendo con las letras CIPP, la esencia del modelo se centra en las cuatro partes de la evaluación tratando de responder a las siguientes preguntas: ¿Qué se necesita hacer? ¿Cómo debe ser hecho? ¿Está siendo realizado? ¿Tuvo éxito en términos de efectividad, impacto, sostenibilidad y transferencia? La lista de control para la evaluación del Modelo CIPP está disponible en www.wmich.edu/evalctr (Sttuflebeam & Shinkfield, 2007:325).

La lista de control está integrada por diez componentes. El primero, relacionado con los acuerdos contractuales para realizar la evaluación, es seguido por los componentes de la evaluación de contexto, entrada, procesos, impacto, efectividad, sostenibilidad, y transferencia. Y los dos restantes tienen que ver con la meta-evaluación y el informe de síntesis final. Los siete componentes centrales de la evaluación se deben utilizar selectivamente y a veces simultáneamente. Los evaluadores deben ser sensibles a atender la información aportada por los implicados en los programas, y la aportada por las fuentes principales de información.

La meta-evaluación debe ser realizada durante todo el proceso. Los evaluadores deben trabajar de modo independiente, sin dejarse influir por cuestiones ajenas a la evaluación. Al final de su trabajo los evaluadores son los testigos del grado en que los estándares de realización a conseguir por esa organización han sido logrados. El componente último de la lista de control de la evaluación proporciona orientaciones para realizar el informe sumativo final, teniendo en cuenta los informes de evaluación formativos obtenidos a lo largo de todo el proceso. (Stufflebeam, 2002: 2).

La comunicación a tiempo de los resultados de la evaluación a quienes han contratado la evaluación y a todas las audiciencias que tengan derecho a estar informadas es un tema clave en la evaluación. En las reuniones ocasionales para mostrar el proceso de la evaluación se deben presentar los informes de un modo claro, ayudándose con programas como el power point, y entregando los informes por escrito para facilitar la diseminación de los hallazgos y el intercambio de valoraciones.

En el informe final el evaluador debe presentar el resumen de los resultados de la evaluación así como la planificación de los siguientes pasos a dar, los acuerdos para una ayuda continuada, especialmente para

la recogida de información, y los cambios necesarios para una evaluación más creíble y útil para el futuro.

Una vez institucionalizada la evaluación, deben mantenerse reuniones, a ser posible una o dos veces al año, para poner en común con todos los interesados en la evaluación los hallazgos que aporta en sus diferentes componentes y llegar a acuerdos sobre los cambios necesarios en la organización y en los proyectos y programas y las ayudas para realizarlos.

El concepto de evaluación que subyace al Modelo CIPP y a su lista de control, es que la evaluación debe valorar e informar a una entidad sobre su mérito, valor y significado con que funciona, teniendo en cuenta las normas del *Joint Committee* (2010), sobre la evaluación que debe ser útil, factible, ética y exacta, requisitos fundamentales para que la evaluación cumpla con su objetivo primordial: la toma de decisiones informada que permita introducir mejoras ajustadas a las necesidades reales que motivaron el diseño del programa y las nuevas necesidades detectadas.

La lista de control para la evaluación se muestra a continuación.

Tabla 2. Síntesis de la lista de control para la evaluación del modelo CIPP

LA LISTA DE CONTROL PARA LA EVALUACIÓN DEL MODELO CIPP	
1. LOS ACUERDOS CONTRACTUALES Las evaluaciones CIPP deben realizarse con acuerdos explícitos previos con los clientes, y deben ser actualizados de acuerdo con las necesidades de cada evaluación.	
Actividades del evaluador	**Actividades de los clientes de la evaluación**
☐ Desarrollar un marco claro y comprensivo del trabajo de evaluación que se va ha hacer.	☐ Aclarar con el evaluador lo que se va a evaluar, su propósito, criterios, y audiencias.
☐ Aclarar la naturaleza, contenidos y plazos para los informes y sesiones formativos.	☐ Asegurar que la planificación de informes es sensible a las necesidades del programa.
☐ Alcanzar acuerdos para proteger la integridad del proceso de información.	☐ Asegurar que el proceso de información será legal, política y éticamente viable.

2. EL CONTEXTO DE LA EVALUACIÓN

Evaluación de las necesidades del contexto, ventajas y problemas dentro del entorno definido.

Actividades del evaluador	Actividades de los clientes de la evaluación
☐ Recopilar y valorar la información previa fundamental, especialmente sobre los beneficiarios y las necesidades a las que se dice debe responder.	☐ Utilizar los hallazgos de la evaluación de cotexto en la selección y clarificación de los beneficiarios potenciales.
☐ Entrevistas con los implicados en el programa para obtener información sobre las necesidades y beneficiarios de la evaluación, así como de los potenciales problemas del programa.	
☐ Valorar las metas del programa a la luz de las necesidades expresas de sus beneficiarios y la utilidad afirmada.	☐ Utilizar los hallazgos de la evaluación, a lo largo del proceso, para valorar la efectividad del programa en satisfacer las necesidades de los beneficiarios.
☐ Discutir los hallazgos de la evaluación en reuniones.	

3. LA EVALUACIÓN DE ENTRADA

Evaluación de las entradas, valoración de las estrategias y de los planes de trabajo y presupuestos.

Actividades del evaluador	Actividades de los clientes de la evaluación
☐ Identificar e investigar si existen programas que puedan servir como modelo para el programa.	☐ Utilizar los hallazgos de la evaluación de entrada para diseñar una estrategia del programa que sea científica, económica, social, política y técnicamente defendible.
☐ Valorar las estrategias propuestas del programa para satisfacer las necesidades valoradas, y su viabilidad.	☐ Utilizar los hallazgos de la evaluación de entrada para asegurar que las estrategias del programa sean viables para satisfacer las necesidades valoradas.
☐ Valorar el mérito de las estrategias del programa comparadas con estrategias alternativas de programas similares.	☐ Utilizar los hallazgos de la evaluación de entrada para formar al personal para realizar el programa.
☐ Finalizar el informe de la evaluación de entrada y proporcionárselo a los clientes.	

4. LA EVALUACIÓN DE PROCESO

Hacer un seguimiento de la evaluación de proceso, documentar y valorar las actividades del programa.

Actividades del evaluador	Actividades de los clientes de la evaluación
☐ Implicar a un grupo en la evaluación para seguir, observar, grabar y proporcionar informes periódicos sobre la aplicación del programa.	☐ Utilizar los hallazgos de la evaluación de proceso como control y fortalecimiento de las actividades del personal. ☐ Y para el fortalecimiento del diseño del programa.
☐ En colaboración con el personal del programa, mantener un seguimiento de los acontecimientos del programa, problemas, costos y asignaciones.	☐ Utilizar los hallazgos de la evaluación de proceso para mantener un seguimiento sobre el progreso del programa.

5. LA EVALUACIÓN DE IMPACTO

Evaluación del impacto que alcanza un programa en relación con las audiencias meta.

Actividades del evaluador	Actividades de los clientes de la evaluación
☐ Implicar al personal del programa, y al grupo de evaluación, para mantener un directorio de personas y grupos a los que sirve, para tener información sobre sus necesidades y los servicios recibidos por el programa.	☐ Utilizar los hallazgos de la evaluación de impacto para asegurar que el programa satisface las necesidades de los beneficiarios a los que debe atender. Y si alcanza necesidades de beneficiarios no previstos.
☐ Evaluar y valorar el grado en que los individuos y grupos que se benefician es consistente con los beneficiarios a los que debe servir.	☐ Asegurar que la planificación de informes es sensible a las necesidades del programa.
☐ Discutir los hallazgos de la evaluación de impacto con los miembros del programa.	☐ Asegurar que el proceso de información sea legal, política y éticamente viable.
☐ Finalizar el informe de la evaluación de impacto, con ayudas visuales para los clientes y contratadores de la evaluación.	

6. LA EVALUACIÓN DE LA EFICACIA Y EFICIENCIA Evaluación de la efectividad del programa y valoración de la calidad y significatividad de sus metas.	
Actividades del evaluador	**Actividades de los clientes de la evaluación**
☐ Entrevistas con los patrocinadores, líderes, personal y beneficiarios del programa y otros grupos interesados para determinar la valoración de logros positivos y negativos.	☐ Utilizar los hallazgos de la evaluación de la efectividad para determinar los efectos positivos y negativos del programa sobre sus beneficiarios.
☐ Realizar estudios de caso en profundidad con grupos de beneficiarios.	☐ Utilizar los hallazgos de la evaluación de la efectividad para descubrir y valorar efectos importantes colaterales.
☐ Implicar al grupo de evaluación y personal del programa para aportar documentación para identificar y confirmar el rango, profundidad, calidad y significado de los efectos del programa sobre los beneficiarios.	☐ Utilizar los hallazgos de la evaluación de la efectividad para examinar si el programa, su planificación y actividades deben ser cambiados.
7. LA EVALUACIÓN DE LA SOSTENIBILIDAD Valorar el grado por el cual las contribuciones del programa están exitosamente institucionalizadas y deben continuar a lo largo del tiempo.	
Actividades del evaluador	**Actividades de los clientes de la evaluación**
☐ Entrevistas con los líderes y personal para identificar sus valoraciones sobre los éxitos del programa que deben ser mantenidos.	☐ Utilizar los hallazgos de la evaluación de la sostenibilidad si el personal y los beneficiarios favorecen la continuación del programa.
☐ En colaboración con el personal, seguimiento de los acontecimientos, problemas, costos y asignaciones.	☐ Utilizar los hallazgos de la evaluación de sostenibilidad como garantía para que las metas y el plan continúen sus actividades.
☐ Entrevista con los beneficiarios para identificar su comprensión y valoración de las provisiones para la continuación del programa.	☐ Utilizar los hallazgos de la evaluación de sostenibilidad como garantía para que el plan y los presupuestos necesarios continúen.
☐ Recoger hallazgos sobre la sostenibilidad del programa de los informes de progreso y final.	

8. LA EVALUACIÓN DE LA TRANSFERIBILIDAD

Valora el grado en que un programa puede tener éxito adaptado en otro lugar.

Actividades del evaluador	Actividades de los clientes de la evaluación
☐ Investigación con muestras representativas de los posibles grupos que puedan adoptar el programa. Cuestiones sobre: 1) Descripción del programa e informe de los hallazgos de la evaluación. 2) Juicio o valoración crítica sobre la relevancia del programa para su situación. 3) Juicio sobre la calidad del programa. 4) Informe sobre si se está aplicando todo o parte del programa.	☐ Uso de los resultados de la evaluación para valorar la diseminación de la información sobre el programa. ☐ Uso de los hallazgos de la evaluación del programa para adaptarlos a otras situaciones.
☐ Visitas y valoración de la adaptación del programa.	
☐ Informes y discusión sobre los hallazgos de la evaluación del programa que pueden ser aplicados a otras situaciones.	

9. LA META-EVALUACIÓN

Es la valoración de la evaluación, para comprobar si cumple con los estándares que deben guiar las evaluaciones de programas.

Actividades del evaluador	Actividades de los clientes de la evaluación
☐ Se han llegado a acuerdos con los clientes de la evaluación sobre que ésta se guiará con las normas del Joint Committee de utilidad, viabilidad, honradez y precisión, o cualquier otro tipo de normas o principios.	☐ Revisar las normas del JC y lograr acuerdos con los evaluadores sobre qué principios seguir para guiar y valorar el trabajo de evaluación. ☐ Considerar una valoración independiente de la evaluación.
☐ Animar y apoyar al cliente para que obtenga una valoración independiente del plan de evaluación, del proceso y/o del informe de resultados.	☐ Mantener un archivo de información pertinente para valorar los acuerdos en contraste con las normas y principios que deben guiar las evaluaciones.

☐ Documentar los procesos de evaluación y los resultados de modo que puedan ser rigurosamente estudiados y evaluados.	☐ Aportar la información y asesorar adecuadamente para legitimar los esfuerzos de evaluar la evaluación.
☐ Utilizar la meta-evaluación periódicamente para asegurar que la evaluación se está realizando adecuadamente.	☐ Tener en cuenta los resultados de la meta-evaluación para valorar cómo se están aplicando los resultados de las evaluaciones.
☐ Valorar y proporcionar comentarios escritos sobre el grado de concordancia que encuentra la evaluación con los estándares y/o los principios que orientan la evaluación, e incluir los resultados en el informe final de la evaluación, a modo de apéndice o anexo.	☐ Considerar incluir un apéndice al informe de evaluación final con reacciones sobre la evaluación, informes de los evaluadores sobre el grado de concordancia entre los estándares y la evaluación realizada, a modo de resultados de una meta-evaluación y documentar los usos significativos de los hallazgos de la evaluación.

10. INFORME FINAL DE SÍNTESIS

Para incluir todos los resultados de la evaluación e informar a todas las audiciencias sobre lo que se pretendió hacer, lo que se hizo, y cómo, qué hemos aprendido, y concretar la línea base de la evaluación útil para sucesivas evaluaciones.

Actividades del evaluador	**Actividades de los clientes de la evaluación**
☐ Organizar el informe para atender a las diferentes necesidades de las diferentes audiencias, incluyendo los antecedentes del programa, su implementación, y sus resultados.	☐ Contribuir a asegurar que los contenidos del informe planeado sea atractivo y sea utilizable por el amplio rango de audiciencias beneficiarias.
☐ Continuando con el ejemplo, en el informe de los antecedentes del programa, incluir secciones discretas sobre la organización que promueve el programa, el origen del programa que se evalúa y el entorno del programa	☐ Ayudar a asegurar que la narración histórica presentada en los antecedentes del programa sea exacta, suficiente, y de interés para su uso, al menos para alguna de las audiencias del informe.

☐ Incluir en el informe de la implementación del programa, secciones que den detalles de los principales componentes del programa planeados, fundamentados, informado el personal, realizado de modo que los grupos de interés en exportar el programa puedan ver como podrían ser realizadas las actividades del programa. Estas secciones serían principalmente descriptivas para utilizarlas con las precauciones pertinentes.	☐ Ayudar a asegurar que la narración sobre la implementación del programa sea exacta, suficiente detallada, para ayudar a otros a comprenderlo así como las posibilidades y los procedimientos de su aplicación (teniendo en cuenta las precauciones pertinentes).
☐ Incluir en el informe de los resultados del programa, los hallazgos de la evaluación (de contexto, entrada, proceso, impacto efectividad, sostenibilidad y trasferencia) y de las conclusiones de la evaluación (fortalezas, debilidades, lecciones aprendidas y línea base de la valoración del mérito, valor y significatividad del programa). Contraste entre las contribuciones del programa con lo que pretendía, lo que los beneficiarios necesitaban y los costos que han supuesto. Y comparación con programas similares.	☐ Informe de la utilidad de los resultados del programa, fundamentando lo que se ha realizado, los fallos, y los esfuerzos comparado con programas similares, y qué lecciones deben ser tenidas en cuenta para futuros programas.
☐ Ilustrar el informe con fotografías, gráficos, citas, comentarios reales. Reconocimientos, apéndices técnicos incluyendo los instrumentos de recogida de datos, el manual seguido por el evaluador, y cualquier otra cosa de interés.	☐ Utilización del informe completo como medio de preservar la memoria de la institucional del programa e informar a las partes interesadas en la empresa.

7. TABLA RESUMEN DEL MODELO CIPP DE EVALUACIÓN DE PROGRAMAS DE STUFFLEBEAM

Tabla 3. Síntesis del enfoque evaluativo de Stufflebeam

LA EVALUACIÓN ORIENTADA A LA ORGANIZACIÓN Y A LA TOMA DE DECISIONES DE STUFFLEBEAM	
Elementos para el análisis	**Características**
Concepciones evaluativas	Estructura básica del modelo de evaluación CIPP.
Definición de evaluación	Proceso de identificar, obtener y proporcionar información útil y descriptiva acerca del valor y mérito de las metas, la planificación, la realización y el impacto de un objeto determinado, con el fin de servir de guía para la toma de decisiones, solucionar los problemas de responsabilidad y promover la comprensión de los fenómenos implicados».
Funciones de la evaluación	Formativa y sumativa
Para quién se hace la evaluación	Para el que toma las decisiones y para todos los miembros que participan en el programa.
Características del evaluador	Especialista al servicio de la organización y de los miembros del programa.
Ámbitos y contextos más adecuados para aplicación	Evaluación de proyectos, programas, personal, instituciones y sistemas.
Fases principales de aplicación	Evaluación de contexto. Evaluación de entrada. Evaluación de proceso. Evaluación de producto.
Metodología evaluativa	Diversa para cada tipo de evaluación. Cuantitativa y cualitativa.

Información a recoger	De las necesidades, los recursos, el funcionamiento y el impacto.
Métodos	Diversos para cada tipo de evaluación. Lista de control para la evaluación.
Instrumentos y técnicas	Diversos: observación, entrevista, cuestionario…
Tipo de informe	Frecuente y para proporcionar una guía para la mejora del programa.
Posible generalización	La que permitan el muestreo y la metodología utilizada.
Utilidad de los hallazgos	Para tomar decisiones de mejora.
Criterios para juzgar las evaluaciones	Utilidad, factibilidad, ética, exactitud y responsabilidad.
Contribuciones	Evaluación de los contextos y de los procesos.

EJERCICIOS DE AUTOEVALUACIÓN

1. Haga una valoración de los problemas que detectaron los miembros del Comité PDK para mejorar la evaluación de programas, y comente en qué medida están solucionados en la actualidad.
2. Concepto de evaluación de programas de Stufflebeam y Shinkfield y su valoración.
3. Haga una tabla de doble entrada en la que recoja los objetivos, propósitos, metodologías y decisiones a las que sirve cada una de las categorías del modelo evaluativo CIPP.
4. ¿Cómo aplicaría el modelo CIPP a la evaluación de un programa educativo? Elija un programa con el que esté familiarizado. Sea preciso,

concreto y breve en su respuesta y reflexione sobre la relación entre lo estudiado en el tema y su aplicación a la evaluación de un programa real.

5. Evalúe un programa que conozca utilizando la lista de control para la evaluación del modelo CIPP y valore la aplicabilidad de la lista al programa.

REFERENCIAS BIBLIOGRÁFICAS

Alkin, M. C. (1969). «Evaluation theory development», *EvaLuation Comment,* 2, 1.

JCSEE (1981). *Standards for evaluation of educational programs, projects and materials*. New York: McGrawHill.

— (1994). *The Program Evaluation Standards*. Thousand Oaks, CA: Sage Publications, Inc.

Normas de evaluación para programas, proyectos y material educativo (1994). México. Trillas.

RAE (2016). *Diccionario de la Real Academia de la Lengua Española*. Edición del Tricentenario. Consultado el 29 de marzo del 2016 en

http://dle.rae.es/?id=QmrGvNl

Stufflebeam, D. L. (1971). «The relevance of the CIPP evaluation model for educational accountability». *Journal of Research and Development in Education,* 5, 1, (1925).

Stufflebeam, D. L. (2002). «CIPP Evaluation Model checklist». *Evaluation checklists project*. www.wmich.edu.

Stufflebeam, D. L. *et al.* (1971). *Educational evaluation and decision making*. Itasca, ILL: Peacock.

Stufflebeam, D. L. & Shinkfield, A. J. (1981). «Meta-Evaluation: Concepts, Standards and Uses». En A. R. Berk. (ed.). *Educational evaluation methodology: The state of the art. Baltimore: The Johns Hopkins University Press.*

— (1987). «Evaluación sistemática. Guía teórica y práctica». Cap. 6: *La evaluación orientada hacia el perfeccionamiento.* Barcelona, Buenos Aires: Paidós-MEC.

— (2007). *Evaluation. Theory, Models and Applications.* San Francisco, CA: John Wiley and Sons. Inc.

Tema 4

Ramón Pérez Juste: La evaluación de programas educativos

1. Objetivos y resultados de aprendizaje.
2. Introducción.
3. La teoría evaluativa.
   - 3.1. Educación y evaluación.
   - 3.2. La evaluación integral, integrada e integradora.
   - 3.3. El concepto de evaluación de programas.
   - 3.4. Características de los programas educativos.
   - 3.5. Función esencial de la evaluación: la mejora.
4. El modelo evaluativo para la práctica evaluativa.
   - 4.1. Etapa inicial: Evaluación del programa en sí mismo.
     - 4.1.1. La calidad intrínseca del programa.
     - 4.1.2. La adecuación al contexto.
     - 4.1.3. Adecuación a la situación de partida.
     - 4.1.4. Criterios de evaluación.
   - 4.2. Etapa procesual: Evaluación del desarrollo de los programas.
     - 4.2.1. Criterios para la evaluación de la implementación de los programas.
     - 4.2.2. Metodología adecuada a la evaluación de procesos.
   - 4.3. Etapa final: Evaluación de los resultados de los programas.
     - 4.3.1. El programa en sus logros.
     - 4.3.2. Criterios y referencias de la evaluación de resultados de la aplicación del programa.
     - 4.3.3. La valoración de los resultados.
   - 4.4. Institucionalización de la evaluación de programas.

5. La pauta para la evaluación de programas educativos.
6. Tabla resumen del modelo de evaluación de programas de Pérez Juste.

Ejercicios de autoevaluación.

Referencias bibliográficas.

1. OBJETIVOS Y RESULTADOS DE APRENDIZAJE

El objetivo principal de este tema es valorar las contribuciones de Ramón Pérez Juste a la Evaluación de Programas, especialmente por su aportación sobre la teoría de programas educativos, y su concepción de la evaluación al servicio de los programas educativos y de las personas relacionadas con ellos junto con su defensa de la complementariedad metodológica para la evaluación de programas.

Se espera que el alumno, al terminar el estudio del tema, consiga los siguientes objetivos de aprendizaje:

— Comprender y valorar la concepción evaluativa de Pérez Juste.

— Comprender las características de los programas educativos.

— Conocer y valorar las etapas para la realización de evaluaciones de programas de acuerdo con el Modelo de Pérez Juste.

— Conocer y valorar los criterios y las referencias para la evaluación de programas del modelo de Pérez Juste.

— Aplicar el enfoque de evaluación de programas de Pérez Juste a la evaluación de un programa educativo.

— Valorar la pauta para la evaluación de programas educativos.

2. INTRODUCCIÓN

Ramón Pérez Juste ha sido catedrático de la Facultad de Educación de la Universidad Nacional de Educación a Distancia (UNED). Ha sido presidente de la sección de Educación de la Asociación Española de

Calidad durante el primer quinquenio del siglo XXI y presidente de la Sociedad Española de Pedagogía. El profesor Pérez Juste falleció en enero del 2017. Descanse en paz.

Sus trabajos más significativos han girado en torno a la Pedagogía Experimental, la Pedagogía Diferencial, la Medida en Educación, la Evaluación de Programas y la Calidad de la Educación. Ha publicado numerosos libros y artículos sobre las temáticas mencionadas, siendo especialmente relevante para esta asignatura su libro de *Evaluación de programas educativos,* del 2006, en el que desarrolla su propio modelo de evaluación de programas educativos.

Su formación pedagógica y metodológica, así como el contacto continuo con la práctica educativa, han hecho posible la elaboración del Modelo de evaluación de programas educativos que a continuación presentamos. Su modelo evaluativo reúne una gran coherencia en su fundamentación teórica y metodológica, y supone una aportación importante a la teoría evaluativa de programas y a la orientación de la práctica evaluativa.

3. LA TEORÍA EVALUATIVA DE PROGRAMAS

Para Pérez Juste, la evaluación de programas es una actividad pedagógica orientada a la mejora de los programas, del educando y del educador de una manera directa, e indirectamente también actúa en la mejora de variables organizativas, técnicas y medioambientales, sobre los programas mismos.

3.1. Educación y evaluación

La educación frecuentemente se considera como el resultado de un proceso sistemático e intencional al servicio de unos objetivos que confluyen en una finalidad general denominada educación integral. Una educación integral trasciende concepciones reductivas de la educación, y abarca el perfeccionamiento y la mejora de todas las dimensiones específicamente humanas para que las personas puedan afrontar los desafíos

en los ámbitos personales, sociales y profesionales, con las mejores posibilidades de éxito (Pérez Juste, 2006, pp. 22-23).

Un proceso educativo así concebido es complejo y duradero en el tiempo, por lo que en él deben confluir actuaciones educativas de diversa naturaleza, desde la evaluación de necesidades para el acercamiento a la realidad de los sujetos, a la planeación de planes, proyectos y programas, al diseño e implementación de los mismos y a la valoración, evaluación de todo el proceso y de sus resultados, necesitando para ello financiación y recursos para su aplicación junto con profesionales, personal con la adecuada preparación. La evaluación se integra como un componente más del programa.

3.2. La evaluación integral, integrada e integradora

La evaluación de programas y de sus resultados en sus beneficiarios no puede permanecer aislada de los contextos en los que se aplican los programas, en las organizaciones, en las instituciones educativas, e incluso, pertenecientes a sistemas más amplios como son los sistemas educativos. Por ello, la evaluación es adjetivada como *integral, integrada e integradora*.

La evaluación de programas es *integral* al considerarla al servicio de objetivos educativos de calidad para la formación integral de las personas.

Es *integrada*, en sintonía con todas las actividades educativas de diagnóstico, planeación, realización, evaluación de los aprendizajes, de los recursos, de los medios, de las personas, al servicio del desarrollo de los programas educativos orientados a conseguir fines educativos: el desarrollo integral de las personas.

Y es *integradora*, en su papel activo sobre la educación, al informar de la calidad de lo que acontece en la educación, informando, orientando sobre aquello en lo que hay que mejorar, en lo que hay que desechar, en lo que hay que innovar, incluyendo los programas, la organización y sus recursos, las personas y los sistemas educativos.

3.3. El concepto de evaluación de programas

Lo dicho en el anterior epígrafe conduce a considerar *su concepción evaluativa*. Pérez Juste define *la evaluación de programas* como:

> El proceso sistemático, diseñado intencional y técnicamente, de recogida de información rigurosa -valiosa, válida y fiable-, orientado a valorar la calidad y los logros de un programa, como base para la posterior toma de decisiones de mejora tanto del programa como del personal implicado y, de modo indirecto, del cuerpo social en que se encuentra inmerso. (1991a:5).

Analiza cada uno de los elementos de su definición de evaluación de programas, del modo siguiente:

> Que sea un proceso representa la exigencia de superar los modelos puramente sumativos, puntuales, desinstitucionalizados, centrados en esencia en la constatación y en el puro control.

Esta consideración concibe la evaluación para la mejora del programa y de sus beneficiarios.

> Que sea sistemático, supone la necesidad de una planificación previa, en la que todas y cada una de las actuaciones y etapas se prevén y tienen su sentido en función de las metas del programa.

Indica un modo de hacer la evaluación, intencionado y justificado.

> Que hablemos de recogida de información parece ponernos ante la necesidad de superar modelos restrictivos de prueba para trascenderlos y alcanzar información de diversa naturaleza, procedente de diferentes fuentes, capaz de contextualizarla, de darle sentido, de facilitar la interpretación; todavía más, con ello queremos significar la atención, la alerta, la sensibilidad a fenómenos colaterales, tales como los efectos no planeados, deseados o no, que pueden llegar a tener repercusión sobre el programa, tanto para modificarlo como para suprimirlo o mantenerlo.
>
> Hablar de que la información sea rigurosa es tanto como hacerlo de un tipo de datos propios de planteamientos científicos; pero al explici-

> tar esas tres notas, queremos dar a entender que sin renunciar a las exigencias fundamentales de la validez, credibilidad y fiabilidad de la información hacemos una apuesta por su riqueza y matización, valía, lo que permitirá siempre interpretación y decisiones más ajustadas a la situación.

Incide en que la información recogida sea amplia, de todos los participantes y de su contexto y veraz.

> Incluir en el modelo la referencia al juicio de valor supondrá la necesidad de explicitar tanto los criterios como las referencias que se habrán de utilizar; pero al incluir las palabras calidad y logros estamos poniendo de manifiesto que nos interesan ambas dimensiones, no contentándonos con planteamientos reduccionistas propios de enfoques meramente economicistas.

Claro está, no centrarse solo en los resultados, sino en si responde a las necesidades y con qué calidad ha respondido a las mismas. Y sigue diciendo:

> Es más, con lo anterior enlaza plenamente, y con notoria fluidez, el hecho de incluir la toma de decisiones; en efecto, la mera constatación de los logros no tiene prácticamente ninguna repercusión directa, ni potencial implícito alguno de mejora, ya que tal información no dice nada sobre las raíces o causas de logros insuficientes, inadecuados o irrelevantes. En consecuencia, será necesario acudir a otros tipos de datos, no estrictamente de resultados —productos—, que orienten sobre qué mejoras deben introducirse, además de identificar tanto el dónde, como el cuándo y el cómo de las mismas, lo que nos lleva a modelos más complejos y más difíciles de articular.

Estos argumentos hacen referencia a uno de los objetivos de la evaluación: para qué se hace, su finalidad y su utilidad.

> Indicar por último, que las decisiones de mejora no se refieren sólo al programa sino, también, al personal, además de poder ser considerado como una redundancia (muchas mejoras del programa no pueden articularse sin que mejore el personal que lo aplica). Quiere ser una manifestación explícita tanto de que el énfasis se sitúa en las personas (el programa es un medio, una herramienta) cuanto una patente mani-

> festación de un enfoque evaluativo holístico, ecológico, donde el programa no puede, o no debe, considerarse al margen de la situación, en forma puramente apendicular, colateral, yuxtapuesta, desintegrada. Más aún, y en relación con el enfoque predominantemente externo a que con anterioridad nos referíamos, esa mejora del personal parece exigir la participación, la colaboración activa de las personas implicadas. (Pérez Juste, 1991).

La definición de Pérez Juste de evaluación de programas trasciende el término que define que, ciertamente, puede resultar limitativo del alcance real que tiene el campo de la evaluación de programas. Se trata de una expresión que engloba la evaluación tanto del programa en su planteamiento como en su desarrollo, proceso, como en sus resultados, teniendo en cuenta el contexto, en su historia y actualidad, el marco en el que se desarrolla, las organizaciones y sus recursos, sus miembros, las personas que lo desarrollan y sus beneficiarios, los alumnos, las familias, la propia organización y su personal y mediante relaciones institucionales con otros profesionales, con otros centros y con la sociedad.

3.4. Características de los programas educativos

Por programa se entiende un documento técnico, elaborado por el profesorado, en el que quedan reflejados tanto sus planteamientos docentes cuanto sus propias intenciones y metas educativas. La dimensión fundamental de programa es que se trata de un plan sistemático de intervención, específica e intencionalmente elaborado, al servicio de metas consideradas valiosas desde una perspectiva pedagógica.

Los programas educativos vienen definidos por las notas siguientes:

— *Complejidad*. Al poner en juego una compleja variedad de elementos organizativos, orientadores, didácticos, adecuados a las necesidades de los sujetos, integrando, de un modo armónico, diferentes saberes multidisciplinares.

Esta complejidad exige la integración de los programas educativos en los Proyectos educativos del centro para lograr la colaboración del personal, lo que redundaría en su eficacia.

— *Dificultad*. Referida a la implementación de los programas por suponer una traducibilidad de conceptos mentales en actividades dirigidas a lograr las metas del programa.

— *Duración*. La consecución de las metas de los programas requiere contar con tiempo suficiente para desarrollar y afianzar los cambios de mejora perseguidos, ya sea de habilidades, actitudes o comportamientos, contando, siempre, con la colaboración de los docentes. Todo ello exige una observación del desarrollo de los procesos.

— *Procesualidad*. Consistente en la traducibilidad de las metas de los programas en la implementación de los mismos, poniendo especial atención a la adecuada interpretación de los fines pretendidos por los programas.

— *Mantenimiento y transferencia*. El verdadero sentido de los programas se orienta a la posibilidad de mantener los logros del mismo y a transferirlos a otros ámbitos y actividades del sujeto, contribuyendo a su formación intelectual y a su madurez y autonomía personal.

3.5. Función esencial de la evaluación: la mejora

Si bien la evaluación es una antigua actividad, el salto cualitativo de la evaluación se reconoce cuando se identifica con la función de mejora. Evaluación para la mejora del objeto evaluado, ya se trate de los programas, las organizaciones, los profesionales, los estudiantes o los objetivos de los programas, desde la evaluación de necesidades en sus contextos, hasta la evaluación de los resultados, pasando por la evaluación del diseño de los programas y de su aplicación, de un modo continuo, formativo, aportando información para la mejora de los evaluados, de sus objetivos, de las personas y de los resultados. La función esencial de la evaluación es la mejora

Tabla 1. Momentos claves para la mejora de los programas a través de su evaluación

FUNCIÓN ESENCIAL DE LA EVALUACIÓN: LA MEJORA	
Momentos	**Elementos**
Previo	Los objetivos: • Del objeto evaluado • Del propio sistema de evaluación
Núcleo de la evaluación	Recogida de información Valoración de la información • Criterios • Referencias Toma de decisiones orientadas a la mejora (Planes)
Posterior	Seguimiento de los planes de mejora

4. EL MODELO EVALUATIVO PARA LA PRÁCTICA EVALUATIVA

La evaluación de un programa se hace con la finalidad de constatar el grado de eficacia con que consigue el logro de las metas para las que fue elaborado. Dada las complejidades que encierra un programa, es asumible que puede mejorarse. Y para poder llevar a cabo las decisiones de mejora cobra sentido la actividad evaluativa. Esta actividad evaluativa está pensada para ser realizada en cuatro momentos, cada uno de ellos con entidad y objetivos propios, y juntos contribuyen a la mejora del programa en su totalidad (Tabla 2. Pérez Juste, 1992:9).

Tabla 2. Momentos para la evaluación del programas y sus componentes

EL MODELO EVALUATIVO PARA LA PRÁCTICA EVALUATIVA		
Momentos	**Dimensiones**	**Objeto**
Inicial: El programa en sí mismo	Calidad intrínseca del programa	• Contenido del programa • Calidad técnica • Evaluabilidad
	Adecuación al contexto	• Respuesta a necesidades carencias • Priorización
	Adecuación a la situación de partida	• Viabilidad

EL MODELO EVALUATIVO PARA LA PRÁCTICA EVALUATIVA		
Momentos	**Dimensiones**	**Objeto**
Procesual: El programa en su desarrollo	Implantación	• Actividades • Secuencias • Tiempo • Flexibilidad
	Marco	• Clima • Coherencia
Final: El programa en sus resultados	Medida y logros	• Constatación • Contraste
	Valoración	• Criterios • Referencias
	Continuidad	• Decisiones • Incorporación de mejoras
Institucionalización de la evaluación	La evaluación como parte de los programas	• Elaboración de planes de mejora • Plan de seguimiento

4.1. Etapa inicial: Evaluación del programa en sí mismo

Se trata de la actividad evaluativa más importante tanto por ser la primera y base de todas las demás como por su contenido —abarca la realidad total del programa— y sobre todo, por sus grandes aportaciones a la mejora y optimización del programa. La evaluación inicial del programa, antes de su puesta en funcionamiento, es esencial por su carácter anticipador, tiene una finalidad eminentemente preventiva.

La finalidad que persigue es establecer la calidad técnica del programa, su viabilidad práctica y su evaluabilidad. Los objetivos que pretende la evaluación inicial son: 1) poner en marcha el programa en condiciones favorables y óptimas, minimizando el riesgo de fracaso y 2) poder excluir o retirar, en caso de fracaso, el programa mismo.

La función evaluativa es formativa, para poder tomar por anticipado las decisiones de mejora que puedan elevar las potencialidades del programa. La información a recoger será sobre el programa: su fundamentación, su formulación y su relación con las necesidades, carencias,

demandas y expectativas de los destinatarios. En suma, se trata de introducir cuantas modificaciones se estimen adecuadas como fruto de la revisión exhaustiva y por expertos del programa.

Entre las dimensiones de la evaluación del programa en sí mismo están 1) la calidad intrínseca del programa, 2) la adecuación al contexto y 3) la adecuación a la situación de partida.

4.1.1. La calidad intrínseca del programa

Para la evaluación de la calidad intrínseca del programa, hay que considerar tanto la entidad de sus metas cuanto la calidad técnica que contribuye al logro, de una manera eficaz, de las metas pretendidas. La calidad intrínseca del programa es concretada por Pérez Juste para su evaluación (1992:11 y ss.), en los siguientes elementos: contenido del programa, calidad técnica y evaluabilidad.

a) Para el análisis del contenido del programa, el criterio fundamental es la coherencia entre el programa y sus bases psico-socio-pedagógicas. Utilizando la contrastación científica, por lo que Pérez Juste afirma que un programa debe convalidarse teorética y experimentalmente, lo que garantizaría su funcionamiento sirviendo a aquello para lo que se ha diseñado, cumpliendo las exigencias de rigor tanto en su fundamentación como en las previsiones de su desarrollo y en las expectativas razonables de éxito.

   La metodología más adecuada sería el análisis crítico del contenido y su contraste con la instrumentación propia del programa. El juicio de expertos también es especialmente adecuado para este tipo de evaluación.

b) La calidad técnica del programa, de otro lado, se apoya en la especificación precisa y clara de las metas, en su congruencia externa entre metas y teoría, en su coherencia interna, entre las partes del programa y en su adecuación a los sujetos y situaciones.

   La metodología adecuada a esta finalidad sería, de nuevo, el análisis crítico del contenido realizado por expertos, referido al diseño del programa.

c) La evaluabilidad del programa se refiere a las posibilidades que el programa contiene para facilitar la información necesaria para decidir sobre su eficacia e incluso identificar los posibles elementos de mejora. Ello supone que el programa debe cumplir ciertas condiciones de calidad, tanto en lo referido a la claridad y precisión del lenguaje que utilice, cuanto a su contenido, explicitando el marco teórico de referencia, las necesidades a satisfacer, los indicadores de progreso y su temporalización, y a su metodología, referida a la especificación de las evidencias para decidir sobre su calidad, a la operativización de las variables en que se concretan las metas y objetivos, en la previsión de los instrumentos adecuados par la recogida de datos, con indicación de sus características técnicas y la previsión del tratamiento y análisis de los datos recogidos, por mencionar algunos de los aspectos descritos por Pérez Juste (1992:13-14).

4.1.2. Adecuación al contexto

La adecuación del programa al contexto supone consideraciones tanto físicas como psicológicas y sociológicas en respuesta a las necesidades sentidas y definidas. El éxito del programa para atender unas necesidades depende del modo de entender y satisfacer las necesidades y expectativas actuales y futuras de los beneficiarios y usuarios finales actuales y potenciales del programa.

En relación con el tema de las necesidades, Pérez Juste ofrece un análisis del término y sus enfoques pertinentes a la teoría de programas, a su evaluación y a los criterios metodológicos. La evaluación de necesidades, de acuerdo con Pérez Juste, conviene afrontarla como cualquier otro tipo de evaluación, como una actividad sistemática, con una planificación, definiendo qué se entiende por necesidad en el contexto de nuestra intervención y establecer el tipo de necesidades que serán el objeto de la evaluación. Después se debe establecer un diseño de la evaluación, concretando su metodología general, sus técnicas e instrumentos a utilizar, los criterios y referencias para su valoración, y los momentos en los que se recogerá la información, la recogida de información y su posterior tratamiento y análisis, valoración, puntos fuertes y débiles, establecimiento de prioridades y toma de decisiones. (Pérez Juste, 2006:169 y ss.).

En cuanto a la valoración de la calidad de la evaluación de necesidades, los criterios a considerar son los siguientes:

— Su referencia y subordinación a la realidad a la que sirve. En nuestro caso, de intervención en educación, la información recogida en la evaluación de necesidades debe atender a la mejora de las personas, de los grupos, de los programas y de las organizaciones.

— La corrección y adecuación de su planificación y diseño a la naturaleza de las necesidades a evaluación.

— Su ejecución eficiente.

— La calidad de la información recogida, siguiendo los criterios de fiabilidad, valides y su carácter valioso par ala toma de decisiones eficiente y eficaz.

— La utilidad y relevancia de la información aportada para los distintos grupos de interés, diseñadores y responsables de programas, aplicadores, evaluadores y beneficiarios.

4.1.3. Adecuación a la situación de partida

Un programa técnicamente bien concebido debe entenderse también como factible, realizable, adecuado a las personas, a sus necesidades y a la situación, viable y realista. La metodología adecuada exige el contraste entre los recursos y las exigencias del programa, la consulta del personal mediante entrevistas, la observación de los comportamientos y los documentos programáticos de los Centros Educativos en donde se implementa el programa.

4.1.4. Criterios de evaluación

Entre otros, para la evaluación de los programas, mencionamos los siguientes criterios:

— Calidad, pertinencia y realismo de objetivos y metas y adecuación de estos a los destinatarios y al contexto. Adecuación a las necesidades.

— Suficiencia de los apoyos, medios y recursos para conseguir los objetivos, en particular sobre la formación, implicación y compromiso de los agentes y demás implicados.

— Calidad técnica de los planteamientos de evaluación.

— Calidad técnica del programa: coherencia con las bases teóricas y con las necesidades a las que trata de responder; congruencia interna entre sus componentes.

— Viabilidad y factibilidad de su aplicación del programa.

— Evaluabilidad: relevancia, suficiencia, claridad y accesibilidad de la información disponible sobre el programa y de la necesaria en los diversos momentos o etapas de su aplicación y evaluación. Análisis de las dificultades detectadas y previsibles para evaluar el programa.

4.2. Etapa procesual: Evaluación del desarrollo del programa

Las características definitorias de los programas son la base argumental para la evaluación de los procesos en su vertiente de evaluación formativa. Dada la larga duración de los programas y sus previsibles reajustes, se hace necesario contar con indicadores intermedios que permitan constatar el rendimiento del programa en su desarrollo así como correcciones de mejora.

El objetivo fundamental de esta etapa es comprobar cómo se desarrolla el programa. Su finalidad es la toma a tiempo de decisiones de mejora, junto con la acumulación de información para mejoras posteriores. Las principal función en programas ordinarios, es la función formativa. La metodología para la recogida y análisis de información se basa en la complementariedad metodológica. Y la información a recoger, se centra en la relacionada con el desarrollo del programa, sus resultados intermedios y sus efectos no planeados —influencia de acontecimientos coetáneos, positivos y negativos—. Los criterios que utiliza son el cumplimiento de los objetivos, la coherencia institucional, la eficacia parcial y la satisfacción de los implicados. Y sus resultados se orientan a la toma de decisiones para los ajustes parciales que sean necesarios realizar.

En este sentido, el marco (contexto, clima, ambiente) es un elemento fundamental que puede interferir en el desarrollo de programa. Asimismo, se debe atender a los efectos no planeados e identificar las circunstancias en que se producen para que puedan ser recogidos, estimulados o potenciados y, sin duda, constatados para tenerlos en cuenta en la evaluación final del programa.

La evaluación de los procesos encierra, al menos, dos propósitos o finalidades: 1) facilitar la toma a tiempo de decisiones inmediatas de mejora, en su función de evaluación formativa y 2) acumular información para posteriores decisiones una vez validado el programa, en su función de evaluación sumativa.

Por lo tanto, la evaluación de los procesos debe atender a las actividades y a la secuenciación de los elementos fundamentales del programa, junto a temporalización y rigidez/flexibilidad con que se aplica y el marco en el que se desarrolla el programa. Cuanta más concreción, en relación con estos aspectos, contenga el programa, más posibilidades tiene de aplicarse de acuerdo con lo que pretende y su evaluación sea coherente y ofrezca resultados realistas, susceptibles de ser tenidos en consideración.

Para la implementación de los nuevos programas se debe contar, necesariamente, con la colaboración del profesorado. Ello conlleva la información y formación explícita del profesorado, antes de poner en acción el nuevo programa, así como una ayuda continua durante el desarrollo de dicho programa.

La necesidad de introducir cambios en el programa será tanto mayor cuanto menos completo y detallado sea el programa. Obviamente, si los objetivos se formulan de un modo vago o amplio, esto lleva consigo un amplio margen de concreción o flexibilidad con el riesgo de que se realicen actividades diferentes para la consecución de los objetivos del programa. Y es necesario tener constancia de estas modificaciones a través de la evaluación de procesos. Junto con lo dicho en el párrafo anterior, hay que tener en cuenta la coherencia institucional como factor potenciador de la eficacia, así como la integración efectiva del programa en el proyecto educativo del centro, manifestada a través de su *normalización*,

acogimiento, valoración por parte del personal..., que pueden considerarse indicadores de un correcto desarrollo.

4.2.1. Criterios para la evaluación de la implementación de los programas

Dos criterios fundamentales señala Pérez Juste (1992:24), para la evaluación de la aplicación de los programas:

1. La conformidad con la planificación, mediante:
   a) La constatación de la existencia de desfases y desajustes. Eficacia parcial.
   b) La comprobación de resultados parciales.
   c) La observación de la aparición de efectos no planeados y
   d) La incardinación efectiva en el Proyecto educativo del centro. Coherencia institucional.
2. La satisfacción del personal, constatada a través de entrevistas a las personas y a los grupos para conocer su visión del programa.

4.2.2. Metodología adecuada a la evaluación de procesos

En coherencia con la complejidad de los programas y su implementación, se impone el principio de *complementariedad metodológica* lo que permite su contraste mediante la *técnica de la triangulación*.

El enfoque de evaluación considerado en relación con quién debe realizar la evaluación, es *mixto*, desde una perspectiva complementaria entre la *evaluación interna y externa*; la evaluación puede ser realizada por el personal propio del programa —evaluación interna— con asesoramiento y apoyo técnico externo —evaluación externa—, salvo que cuente con personal adecuado para cumplir esta misión. De ese modo se potencia la participación, se aumenta la reflexión y la racionalidad y se logran mayores niveles de implicación del personal, lo que beneficia tanto a los profesionales como al logro de las metas del programa. O desde una perspectiva externa, mediante la contratación de un evaluador externo, que deberá, además

de recoger la información necesaria para evaluar el programa, informarse a través de la consulta a todos los grupos implicados en el programa.

Complementariedad de los métodos. El reconocimiento explícito de que es el objetivo de investigación el que determina la opción metodológica (Husén, 1988, en Pérez Juste, 1992:2), junto con la propuesta de Merton (1989) del *eclecticismo disciplinado*, lleva a optar por una complementariedad de métodos sucesiva, en donde la metodología cualitativa ayuda a un conocimiento en profundidad de las situaciones, además de sugerir variables, enfoques y criterios para la planificación metodológica y la construcción de los instrumentos, la comprensión de los resultados y la identificación de los elementos en los cuales concentrar los esfuerzos de mejora de los programas.

Complementariedad de técnicas. Acorde con la complementariedad metodológica está la recogida de información utilizando diferentes técnicas. Recomienda el uso de la observación como técnica básica y esencial para el análisis y valoración de los procesos; el diálogo y las entrevistas formales con el personal implicado en el programa, utilizado como técnica complementaria a la observación que puede añadirnos mayor comprensión acerca de lo observado, además de contraste de información; técnicas formales de recogida de información tales como los autoinformes, los cuestionarios, inventarios, escalas, etc., que suponen alternativas complementarias y de contraste, mediante la técnica de triangulación, con la información recogida mediante la observación y la entrevista. La utilización de pruebas para comprobar el logro de los objetivos en diferentes etapas, sería necesaria en algunas ocasiones.

Organización y valoración de los datos. La existencia de datos procedentes de diversas fuentes hace necesario acudir al contraste de resultados mediante procesos de triangulación; el juicio de expertos como el utilizado en el enfoque de auditoría, en el que el experto analiza las diferentes fuentes de información y la calidad de los datos acudiendo, si fuera necesario a pruebas de detalle, aportan, en la valoración, un nivel aceptable de evidencia (Martínez Aragón, 1989, en Pérez Juste, 1992:32).

En la Tabla 3 podemos ver la metodología para la evaluación de programas, de acuerdo con el enfoque de Pérez Juste.

Tabla 3. Planteamiento metodológico de la evaluación de los procesos

Metodología	
Objetivo	Acumular información para tomar decisiones de mejora, una vez desarrollado y evaluado el programa.
Objeto	Implementación del programa: • Actividades. • Temporalización. • Secuencias. • Flexibilidad.
	Marco: • Clima social. • Coexistencia con otros programas. • Coherencia institucional.
Criterios	Conformidad con la planificación: • Desfases/desajustes. • Resultados parciales. • Efectos no planeados. • Incardinación en el Proyecto educativo. • Satisfacción del personal.
Planteamiento metodológico	Principio de complementariedad: • En los evaluadores: Internos y externos. • En los métodos: Cuantitativos y cualitativos. Complementariedad sucesiva. • En las técnicas: La observación como técnica básica y esencial. El diálogo y las entrevistas formales. Técnicas formales de recogida de información.
Incardinación de los resultados en el marco general	El enfoque de auditoría: • Control interno. • Credibilidad de los datos. • Pruebas de cumplimiento y sustantivas de detalle.

4.3. Etapa final: Evaluación de los resultados de los programas

El análisis evaluativo de los procesos, en el marco específico de la evaluación de programas, supone la conveniencia de incardinar los re-

sultados de la evaluación de los procesos con la evaluación de los logros del programa, e, incluso, con la propia evaluación inicial. La finalidad de la evaluación de los resultados es la de comprobar la eficacia del programa. Su función evaluativa es fundamentalmente sumativa, pero orientada a la mejora de los programas en su nueva edición, en función formativa, o sumativa formativa. La metodología a utilizar es diversa: diseños y pruebas diversas; análisis estadísticos adecuados. La información a recoger se centra en los resultados conseguidos junto con los efectos no planeados. Y los criterios son la eficacia, eficiencia, efectividad, satisfacción, impacto, junto con las referencias en relación con el progreso, logro del dominio, superioridad. Las decisiones a las que debe orientar es la conveniencia de mantener, suprimir, mejorar el programa.

4.3.1. El programa en sus logros

El objetivo de la evaluación de los logros de un programa es comprobar la eficacia del mismo, ya sea en sentido absoluto o relativo, de cara a mantenerlo, mejorarlo, cambiarlo o suprimirlo. La metodología a utilizar es diversa e incluye todos los diseños y técnicas clásicos utilizados en la investigación educativa; el evaluador deberá seleccionar aquellos más pertinentes al enfoque evaluativo, a los objetivos pretendidos y a las exigencias propias de los modelos estadísticos (Pérez Juste, 1992:34). Su función es fundamentalmente sumativa, pero también para futuras ediciones del programa es formativa, para la mejora del mismo. La información a recoger será acerca de los resultados en relación con los objetivos, así como los efectos, positivos o negativos, no planeados.

4.3.2. Criterios y referencias de la evaluación de resultados de la aplicación del programa

Entre los criterios consideramos los siguientes:

— Eficacia: comprobar el grado en que los objetivos pretendidos han sido conseguidos. Se mide en términos de resultados, mediante pruebas. Pueden encontrase efectos, resultados, no planeados, por in-

fluencia de otras variables, por ejemplo, por otros programas. No se puede medir en términos de opinión ni en términos de satisfacción.

— Eficiencia: relación entre los recursos utilizados para los resultados conseguidos.

— Efectividad: es la medida en la cual un programa ha alcanzado unos efectos.

— Satisfacción al final del programa, que se añade a la conseguida a lo largo del programa. En relación con las expectativas sobre el programa de los destinatarios, agentes y otro personal interesado/afectado.

— Impacto del programa en el contexto en que se aplica, efectos posteriores a la terminación del programa.

Y entre las referencias consideramos las siguientes:

— Situación de partida: el programa frente a si mismo (progreso).

— Niveles pre-especificados - evaluación criterial (logro del dominio).

— Niveles de otros programas (superioridad).

En cuanto a las referencias, se podrá decir si los resultados conseguidos por este programa han mejorado en relación con los conseguidos por el programa anterior, *evaluación idiosincrática*, en relación con los conseguidos por otros programas, *evaluación normativa*, o en relación con los objetivos del propio programa, el dominio, *evaluación criterial*.

4.3.3. La valoración de los resultados

En torno a la valoración de los resultados, Pérez Juste considera un problema el de su interpretación, dependiendo de los distintos juicios y decisiones, incluso partiendo de los mismos datos. Y esto es especialmente complicado en la evaluación de programas porque requiere de la integración de valoraciones diversas en un único juicio final, de carácter sumativo. Aconseja maximizar la objetivación de los propios juicios (los del evaluador o equipo), explicándolos razonada-

mente, y «dejando al lector, a la vista de los datos y juicios, que tome sus propias decisiones».

4.4. Institucionalización de la evaluación de programas

Con la expresión institucionalización de la evaluación, Pérez Juste (2006:257), quiere señalar que la evaluación de programas debe formar parte del mismo programa, con la finalidad de que el proceso de elaboración del programa, aplicación, evaluación de resultados, decisiones de cambio y mejora, formen parte de un continuo, a modo de proceso de mejora mediante la auto-evaluación, con ciclos sucesivos de evaluación – mejora – evaluación – mejora, siguiendo el ciclo de mejora continua de Deming de Planificar, Realizar, Evaluar y Actuar en consecuencia.

La finalidad de la institucionalización de la evaluación es integrar la evaluación en el programa al servicio de su mejora y de su eficacia, así como de la mejora de la organización y del personal en su contexto educativo. Las funciones que pretenden es la mejora de las organizaciones hacia la excelencia. Y la metodología es la propia de la mejora continua.

En este momento se logra la plena integración del programa y de su evaluación, lo que viene a redundar en la mejora tanto del programa como de la propia metodología de la evaluación. De este modo, el último momento del proceso de evaluación de un programa se convierte en el momento inicial del siguiente ciclo, mediante las modificaciones del programa y la inclusión de planes de mejora, en un proceso de mejora continua en el que la evaluación queda institucionalizada en la organización educativa.

La institucionaliza la evaluación debe considerarse como un elemento más de las dimensiones de la organización de los centros educativos.

5. PAUTA PARA LA EVALUACIÓN DE PROGRAMAS EDUCATIVOS

A continuación, y con el carácter de mera propuesta, ejemplo o documento de ayuda que sirva de guía a los evaluadores, se presenta una *pau-*

ta que incorpora una serie de *indicadores,* esto es, de aspectos en los que fijar la atención por representar manifestaciones de las variables a valorar, a los que deberían aplicarse los oportunos *criterios*. Los elementos componentes de la *Pauta* se presentan en forma de pregunta a la que dar contestación. (Pérez Juste, 2006: 527 y ss.).

La *Pauta* está planteada sobre la base de un programa que se diseña a partir de una evaluación de necesidades; sin embargo, mediante los oportunos cambios es perfectamente aplicable a programas que responden a características y carencias de los destinatarios conocidas de antemano por su generalidad y que, por tanto, no necesitan formalmente de tal evaluación. Recogemos algunos ítems de la pauta.

Tabla 4. Resumen de la pauta para la evaluación de programas educativos.

LA PAUTA PARA LA EVALUACIÓN DE PROGRAMAS EDUCATIVOS	
I. MOMENTO INICIAL: EVALUACIÓN DEL PROGRAMA EN SÍ MISMO	
1. ADECUACIÓN. PERTINENCIA	
Ítem	1.1. A los destinatarios
1.	¿Existen datos en la organización (registros, informes, actas…) sobre las necesidades y carencias de los destinatarios del programa? ¿Se conocen sus demandas y expectativas? ¿Se acude, cuando es necesario, a personas expertas para su identificación y valoración?
3.	¿El programa toma en consideración las necesidades, carencias y demandas detectadas, identificadas y valoradas mediante algún procedimiento sistemático de evaluación de necesidades?
5.	¿Se diseña el programa a partir de la priorización de las necesidades identificadas?
	1.2. Al contexto
7.	La organización del centro educativo, ¿favorece y estimula la participación, la corresponsabilidad y la implicación del personal en su proyecto y objetivos?
9.	¿Se da cooperación con las familias cuando es necesaria para el desarrollo del programa?
10.	¿Se busca / logra la cooperación de agentes sociales externos a la organización cuando son necesarios para el éxito del programa?

	2. CALIDAD INTRÍNSECA DEL PROGRAMA
	2.1. Contenido del Programa
11.	¿Se han explicitado las bases científicas y socio-psicológicas del programa?
13.	¿Están actualizados o desfasados los contenidos?
14.	Los contenidos incluidos, ¿se consideran relevantes desde perspectivas científicas, sociales, psicológicas y pedagógicas (valor formativo)?
	2.2. Calidad Técnica del Programa
15.	Los objetivos del programa, ¿son suficientes y adecuados a las necesidades / carencias /demandas / expectativas que lo justifican?
16.	¿Se incluyen en el programa objetivos, medios —entre ellos, las actividades, la metodología o la evaluación formativa— y sistema de evaluación del propio programa?
20.	¿Se da adecuación del programa a las características diferenciales, motivación, intereses, capacidad, de sus destinatarios?
	2.3. Evaluabilidad
21.	La información contenida en el programa de cara a su posterior evaluación ¿se considera suficiente, relevante, adecuada ?
22.	¿Se dispone de información clara y precisa sobre aspectos metodológicos y de contenido del programa?
23.	La información necesaria para la futura evaluación, ¿es pública? ¿resulta accesible a las partes interesadas e implicadas?
	3. ADECUACIÓN A LAS CIRCUNSTANCIAS: VIABILIDAD
24.	¿Son realistas los objetivos formulados?
26.	¿Existen en el centro los medios y recursos necesarios y suficientes para un correcto desarrollo del programa?
27.	¿Se encuentran capacitados los responsables para el desarrollo de programa? Cuando sea necesario, ¿han recibido formación específica para ello?
29.	¿Se realizan reuniones del equipo de profesores responsables del programa para planificar la implantación e implementación del programa y para asegurar la coherencia de planteamientos?

30.	¿Se da entre el equipo de profesores, y en el centro en general, un trabajo de tipo cooperativo?
32.	¿Apoyan el programa el claustro de profesores, los padres y demás personal implicado / afectado?
	II. SEGUNDO MOMENTO: PROCESO DE APLICACIÓN DEL PROGRAMA
	1. PUESTA EN MARCHA DEL PROGRAMA
33.	¿Se toman las iniciativas pertinentes para que la implantación del programa sea bien acogida?
35.	La metodología utilizada, ¿resulta adecuada para el desarrollo de los objetivos del programa?
36.	¿Muestran los alumnos interés / motivación hacia las actividades del programa?
40.	Los niveles parciales de logro ¿se consideran alejados de, o acomodados a, los propuestos?
42.	Los recursos planeados ¿resultan suficientes y adecuados? ¿Están disponibles?
44.	¿Se cuenta con un sistema de registro de la información que facilite la posterior evaluación y, sobre todo, una mejor gestión de los procesos (organizaciones que aprenden)?
	2. MARCO DE APLIACIÓN DEL PROGRAMA
45.	El ambiente general del centro, ¿es positivo (relaciones de confianza, amables, cordiales entre el personal)? ¿Resulta favorable al programa?
46.	¿Se aprecian conflictos y tensiones entre el personal del centro que puedan repercutir en el desarrollo del programa? En caso afirmativo, ¿se toman las medidas oportunas para su resolución, sin que afecten al desarrollo del programa?
48.	¿Se detecta algún conflicto entre los objetivos y planteamientos del programa con los del proyecto educativo o con los de otros profesores del equipo o del centro?
49.	¿Están previstas las actuaciones para la resolución de estos problemas que, indirectamente, pueden afectar al desarrollo y eficacia del programa? Por ejemplo: ¿está previsto que la dirección de la organización intervenga en la resolución del conflicto, en caso de producirse?
52.	¿Se aprecia satisfacción en los responsables del programa, en sus destinatarios y en las demás partes implicadas e interesadas?

	III. TERCER MOMENTO: EVALUACIÓN FINAL DEL PROGRAMA
	1. MEDIDA Y LOGROS
53.	¿Se recurre a técnicas variadas de recogida de datos, acordes con la diversidad de los objetivos del programa?
55.	Las técnicas e instrumentos utilizados para decidir sobre la eficacia del programa, ¿son adecuadas a las características de los contenidos / objetivos que se desea evaluar?
57.	¿Se planificaron las pruebas, y demás instrumentos elaborados para la recogida de datos, en el momento de diseñar el programa, a fin de asegurar la máxima coherencia con los objetivos?
58.	¿Se especificaron los criterios de calificación y de los niveles de logro del programa?
60.	¿Se trata de identificar posibles efectos no planeados, pero deseables (efectividad)? ¿Y los efectos no planeados pero indeseables (tensiones entre el profesorado, rechazos...)?
61.	¿Hay indicios y datos sobre la eficiencia del programa? ¿Y sobre su impacto en la organización y, si procede, en otros ámbitos próximos del contexto?
	2. VALORACIÓN
62.	¿Se especificaron, de modo claro y preciso, los criterios y las referencias para valorar los resultados?
63.	¿Se aplican los criterios y las referencias de conformidad con las previsiones y con las exigencias técnicas?
64.	Se dispone de información rica y matizada sobre los momentos inicial y procesual del programa como base para valorar los resultados?
65.	¿Se elabora el informe con la participación de las diferentes partes interesadas e implicadas, con diferentes niveles de responsabilidad según las funciones de cada uno?
66.	¿Se recogen en el informe los puntos fuere y débiles del programa?
67.	El informe ¿logra el acuerdo, el consenso, de las partes interesadas? ¿Se ha logrado la implicación del personal en las decisiones y, en particular, en los planes de mejora? ¿Se da a conocer a todas las partes con interés en el tema?
68.	¿Se elaboran planes de acción para reforzar los aspectos positivos y eliminar los negativos?

	3. CONTINUIDAD
71.	¿Existe un proceso institucionalizado de evaluación, de forma que los resultados se plasmen en nuevos procesos programadores?
72.	¿Se asignan tareas concretas a los diferentes responsables de los cambios a introducir?
73.	¿Se da un seguimiento sistemático de las acciones de mejora emprendidas a partir de la evaluación hasta darlas por concluidas?
	IV. CUARTO MOMENTO: INSTITUCIONALIZACIÓN DE LA EVALUACIÓN
74.	En el caso de un programa cuya aplicación debe tener varias ediciones sucesivas, ¿se aprecia un proceso sistemático y continuo en la secuencia evaluación > mejora > evaluación > mejora?

6. TABLA RESUMEN DEL MODELO DE EVALUACIÓN DE PROGRAMAS DE PÉREZ JUSTE

Tabla 6. Síntesis del modelo evaluativo de Pérez Juste.

LA EVALUACIÓN DE PROGRAMAS EDUCATIVOS DE PÉREZ JUSTE	
Elementos para el análisis	**Características**
Concepciones evaluativas	Características de los programas, evaluación del programa, proceso y resultados.
Definición de evaluación	Proceso sistemático de recogida de información orientada a valorar la calidad y los logros para decisiones de mejora del personal y del programa.
Funciones de la evaluación	Dependiendo del momento de la evaluación, este modelo cumple una función preventiva, formativa y sumativa. • Primer momento: preventiva. Formativa. • Segundo momento: formativa. • Tercer momento: sumativa o sumativa formativizada.
Para quién se hace la evaluación	Para todos los interesados en el programa.

Elementos para el análisis	Características
Características del evaluador	Experto.
Ámbitos y contextos más adecuados para aplicación	Contextos educativos.
Fases principales de aplicación	Evaluación del programa en sí mismo. Evaluación del desarrollo del programa. Evaluación de los resultados del programa. Institucionalización de la evaluación.
Metodología evaluativa	Cuantitativa y cualitativa.
Información a recoger	Diversa. Del programa, de los contextos, de los procesos y de los resultados.
Métodos	Diversos. Principio de complementariedad metodológica
Instrumentos y técnicas	Variados. Observación, entrevista, cuestionarios, pruebas de rendimiento, Delphi, Juicio de expertos y registro.
Tipo de informe	Descriptivo.
Posible generalización	La que permita el diseño.
Utilidad de los hallazgos	Para la mejora del personal y del programa, y de la comunidad.
Criterios para juzgar las evaluaciones	Rigurosidad (evaluaciones valiosas, válidas y fiables)
Contribuciones	Mejora de los programas del educando y del educador. Indirectamente: organizativas, técnicas y medioambientales. Desarrollo de teorías de programas, complementariedad metodológica y técnica.

EJERCICIOS DE AUTOEVALUACIÓN

1. Comente la definición de evaluación de programas de Pérez Juste.
2. Mencione las características de los programas educativos.
3. Haga una síntesis de las etapas para la realización de la evaluación de un programa de acuerdo con el enfoque de Pérez Juste.
4. Los criterios y las referencias en cada uno de los momentos de la evaluación de programas.
5. Haga una síntesis de los métodos y técnicas aplicadas a cada una de las etapas o fases de la evaluación de programas.
6. ¿Cómo aplicaría el modelo de Pérez Juste a la evaluación de un programa educativo con el que esté familiarizado?
7. Haga una tabla de doble entrada para aplicar el modelo de Pérez Juste a la evaluación del programa que ha seleccionado. Sitúe en las filas cada momento de la evaluación del modelo y en las columnas los objetivos y función de la evaluación, las dimensiones a evaluar, la información a recoger, los criterios y referencias y las técnicas e instrumentos de recogida de información. Trate de ser preciso y reflexione sobre la relación entre los estudiado en el tema y su aplicación a la evaluación del programa educativo que conoce.
8. Analice el concepto de 'evaluación formativa' en la evaluación de programas. Su finalidad, cómo se aplica, principales autores que la han desarrollado, metodologías e instrumentos para realizarla.
9. Principales funciones de la evaluación sumativa, mas allá de lo que encuentre escrito por diversos autores; reflexione sobre lo que puede aportar la evaluación sumativa en la evaluación de programas y en la mejora de la calidad de la educación.
10. Reflexione sobre el concepto de «institucionalización de la evaluación» más allá de lo explicado en el tema.
11. Aplique el enfoque de evaluación de programas de Pérez Juste a la evaluación de un programa educativo, mediante la Pauta de evaluación, y valore su adecuación y utilidad.

REFERENCIAS BIBLIOGRÁFICAS

Martínez Aragón, L. (1989): «El enfoque empresarial (modelo de auditoría) aplicada a la evaluación externa de Centros Educativos». En R. Pérez Juste y L. Martínez Aragón: o.c.

Merton, R. K. (1975). «Structural analysis in sociology». En P. Blau (Ed.): *Approaches to the study of social structure*. New York: The Free Press.

Pérez Juste, R. (1991a.). «Evaluación de programas». Ponencia presentada en el *Primer Simposiun Nacional sobre Programas de Enseñar a Pensar*. Granada.

— (1991b). «La evaluación de los procesos en los programas educativos» *Seminario internacional de Evaluación de programas educativos,* Madrid.

— (Coordinador) (2000). «Evaluación de Programas educativos». En *Revista de Investigación Educativa.* Monográfico. Vol. 8, 2.

— (Coord.) (2000): «La Calidad de la Educación». Madrid. Narcea.

— (2006). *Evaluación de programas educativos*. Madrid: La Muralla.

— (2016). «¿Quo vadis, evaluación? Reflexiones pedagógicas en torno a un tema tan manido como relevante». *Revista de Investigación Educativa,* 34(1), 13-30. DOI: http://dx.doi.org/10.6018/rie.34.1.239381

Pérez Juste, R. y Martínez Aragón, L. (1989). *Evaluación de centros y calidad educativa*. Madrid: Cincel.

Tema 5

Los modelos lógicos como marco teórico y de organización para el diseño, aplicación y evaluación de programas educativos

1. Objetivos y resultados de aprendizaje.
2. Introducción.
3. Los modelos lógicos para el diseño y evaluación de programas.
   - 3.1. Los modelos lógicos como secuencia para atender a unas necesidades.
   - 3.2. Marco teórico y de organización orientado a la consecución de los resultados.
4. La evaluación del programa.
5. Los criterios de calidad.
6. Las metodologías de investigación para la evaluación.
   - 6.1. Las metodologías de investigación mixtas para atender la complejidad.
   - 6.2. La elaboración de instrumentos para la recogida de información.
   - 6.3. Compromiso con el rigor y la responsabilidad en la investigación evaluativa.
7. Elaboración de los informes de la evaluación de programas.
8. Utilización de los resultados de la evaluación.
9. Conclusiones.

Ejercicios de autoevaluación

Referencias bibliográficas.

1. OBJETIVOS Y RESULTADOS DE APRENDIZAJE

Tras el estudio de los temas que preceden a éste, centrados en el concepto, funciones, teorías y modelos de Evaluación de Programas, el estudiante conocerá los fundamentos para poder realizar evaluaciones de programas educativos y sociales reales. Este tema tiene como principal objetivo, profundizar en los procedimientos para evaluar programas. De manera más concreta, se espera que el estudiante logre los siguientes objetivos de aprendizaje:

— Comprender y valorar la finalidad de los modelos lógicos para el diseño y evaluación de programas.

— Comprender y valorar los elementos a tener en cuenta en el diseño de programas y su aplicación, a la luz de los criterios a tener en cuenta para garantizar su éxito.

— Comprender y valorar los criterios de calidad que subyacen en los programas y en su aplicación y que deben guiar su evaluación.

— Comprender y valorar el modo de realizar los informes de la evaluación y la potenciación del uso de sus resultados.

— Planificar y realizar evaluaciones de programas siguiendo las orientaciones de los modelos lógicos.

2. INTRODUCCIÓN

En las últimas tres décadas, los modelos de evaluación de programas con objetivos de mejora y de aseguramiento de la calidad para atender a

las necesidades y problemas educativos y sociales, se han convertido en elementos esenciales para dirigir los esfuerzos por la mejora continua de los programas y de las organizaciones.

Para ayudar a la gestión de los programas, los modelos lógicos se constituyen en una herramienta que contribuye a mostrar un cuadro simplificado de un programa, proyecto, iniciativa o intervención para responder a unas necesidades en un contexto dado, mostrando las relaciones lógicas entre los recursos que se requieren, las actividades que deben realizarse y los beneficios o cambios que deben resultar.

3. LOS MODELOS LÓGICOS PARA EL DISEÑO Y EVALUACIÓN DE PROGRAMAS

3.1. Los modelos lógicos como secuencia para atender a unas necesidades

El modelo lógico, sistémico, es la base para elaborar una secuencia convincente del rendimiento esperado de un programa, diseñado para atender a unas necesidades, siendo sus elementos los recursos, las actividades, los resultados inmediatos en los beneficiarios a los que atiende, junto con los resultados en el corto plazo y su impacto, considerando las influencias externas relevantes (Mclaughlin & Jordan, 1999). El modelo lógico muestra las relaciones entre los diferentes elementos de la secuencia a partir de unas necesidades para mostrar la hoja de ruta hacia los resultados.

Los modelos lógicos son la esencia de la planificación de la evaluación de programas, la gestión de los programas y la comunicación de todo ello a los implicados y que tienen derecho a saber para comprender lo que se pretende realizar para atender a unas necesidades, a unos problemas, con unas prioridades fundamentadas en la teoría y en las experiencias de programas similares para atender situaciones parecidas. Los modelos lógicos ayudan a la descripción de necesidades, a la formulación de prioridades para atenderlas, al diseño y gestión de los programas para conseguir los resultados pretendidos en coherencia con las necesidades y para planificar y realizar su evaluación.

Desde la perspectiva de la evaluación, un modelo lógico atiende al desarrollo de los componentes de la evaluación y a la forma en que se armonizan con los objetivos generales del programa y sus teorías subyacentes, ayudando a priorizar y a enfocar las cuestiones de interés, asegurando la recogida adecuada de los datos de la evaluación, para, tras su análisis, hacer recomendaciones informadas sobre las mejoras que necesita incorporar el programa. (Newton, Poon, Nunes & Stone, 2013, en Freedman *et al.,* 2014: 81).

3.2. Marco teórico y de organización orientado a la consecución de los resultados

Los modelos lógicos ofrecen un marco teórico y de organización para el diseño, desarrollo y comprensión de los programas y de sus objetivos, de su planificación, de su aplicación y de su evaluación, para la comprobación de sus resultados y sus logros, dependiendo de la gestión del programa y de los recursos para su adecuada aplicación.

Todos los elementos deben estar alineados, basados en la teoría y en las evidencias empíricas de resultados de programas similares, lo que ayudará a explicitar las teorías de la acción para el cambio y contribuirá a reconciliar las diferentes expectativas entre los grupos con el fin de lograr acuerdos y apoyos para seguir adelante con el programa. Los modelos lógicos son adecuados para la planificación, diseño, gestión y evaluación del programa, y así mismo, para la comunicación de sus resultados, como clave de su éxito y sostenibilidad.

Los modelos lógicos comienzan con una orientación hacia los resultados que se pretenden conseguir, las metas, que son la razón de los programas en respuesta a unas necesidades valoradas. Después se diseña el programa y se analizan y concretan los recursos necesarios, junto con las estrategias y actividades para la aplicación del programa y la consecución de los resultados inmediatos, a partir de las actividades que realiza el programa, en los participantes y beneficiarios del mismo, junto con los efectos que se quieren conseguir con el programa en el medio y largo plazo. (Figura 1).

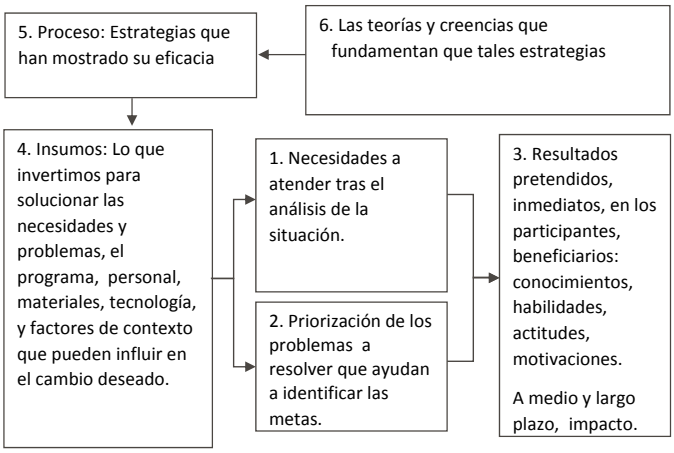
5. Proceso: Estrategias que han mostrado su eficacia
6. Las teorías y creencias que fundamentan que tales estrategias
4. Insumos: Lo que invertimos para solucionar las necesidades y problemas, el programa, personal, materiales, tecnología, y factores de contexto que pueden influir en el cambio deseado.
1. Necesidades a atender tras el análisis de la situación.
2. Priorización de los problemas a resolver que ayudan a identificar las metas.
3. Resultados pretendidos, inmediatos, en los participantes, beneficiarios: conocimientos, habilidades, actitudes, motivaciones.
A medio y largo plazo, impacto.

Figura 1. Planificación del programa orientado a los resultados para atender las necesidades.

La figura 2 muestra el marco de trabajo del modelo lógico.

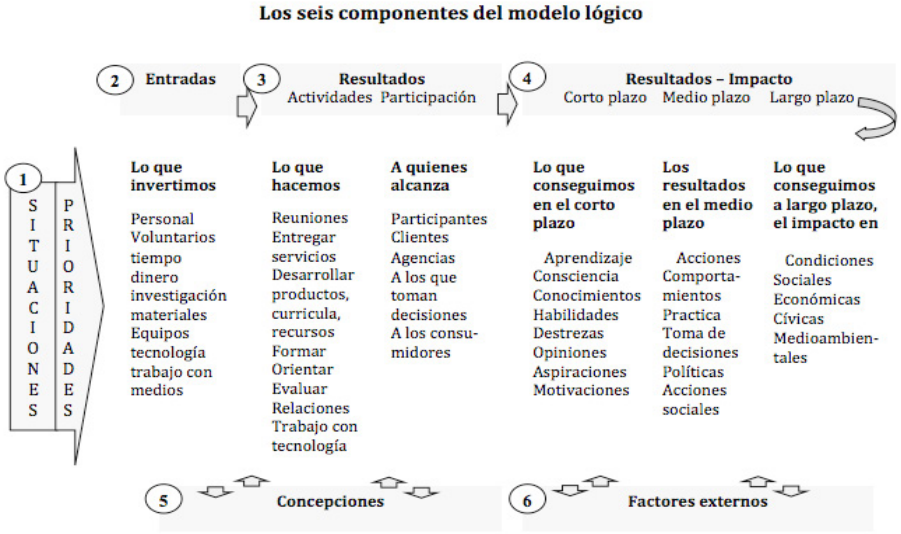

Figura 2. Marco de trabajo del modelo lógico.

Los modelos lógicos proporcionan una representación visual y esquemática del programa en términos de sus insumos, actividades y resultados, mostrando una trayectoria lineal hacia los logros futuros que pretende conseguir. Es una herramienta para la gestión del programa que

ayuda a clarificar las metas y sus relaciones, conseguir acuerdos y consensos, identificar lagunas que necesitan ser subsanadas, explorar enfoques alternativos (Tabla 1).

Tabla 1. Elementos del modelo lógico.

1. Insumos (Recursos)	2. Actividades del programa	3. Resultados inmediatos	4. Resultados a medio plazo	5. Resultados a largo plazo
Informe de la evaluación de necesidades. El programa y su teoría. Equipos, recursos y tecnologías. La organización y la planificación estratégica para la aplicación del programa. Tiempo, dinero. Clima organizativo, de centro y de aula. Profesionales que aplican los programas y personas a las que benefician. Mención de los criterios de evaluación del propio programa, de acuerdo con sus objetivos, destinatarios y recursos. Elaboración de indicadores para la evaluación de los procesos y de los resultados. Determinación de los niveles de logro.	Descripción de cada una de las actividades planificadas necesarias para desarrollar el programa. Estrategias de formación, de motivación, de participación y de colaboración. Desarrollo de materiales y recursos. Evaluación de los procesos, orientación formativa.	Descripción de los resultados pretendidos a los que conduce cada una de las actividades planificadas. Evaluación procesual, continua, formativa y propuestas de mejora. Resultados conseguidos. Identificación de información útil para los ajustes y mejoras del programa y de los procesos.	Comparación entre los resultados conseguidos y los pretendidos por la actividad planificada. Concreción de los niveles de logro. Empoderamiento de la comunidad. Convergencia de estrategias. Evaluación sumativa, e informada por la evaluación procesual. Resultados conseguidos.	Descripción los resultados a largo plazo, y del impacto que se espera conseguir. Evaluación del impacto. Relaciones de colaboración. Responsabilidad individual. Resolución de conflictos adaptados. Respuestas adaptadas a las situaciones. Impacto conseguido. Mejoras en las condiciones sociales, económicas, cívicas, medioa-ambientales
6. Factores claves del contexto que pueden influir sobre su éxito				

Los modelos lógicos se guían por hipótesis que señalan la dirección del cambio que se pretende y asegurar y comprobar que se dan los factores que pueden incidir en él, del siguiente modo:

— «Si los objetivos del programa son realistas y plausibles de conseguir, entonces se conseguirán».

— «Si se cuenta con los recursos necesarios para la aplicación del programa, entonces será más factible que se aplique del modo adecuado para conseguir los objetivos pretendidos».

— «Si los procesos se realizan de modo que alcancen los objetivos pretendidos en toda su extensión, entonces se asegurará la consecución de los mismos» (Figura 3).

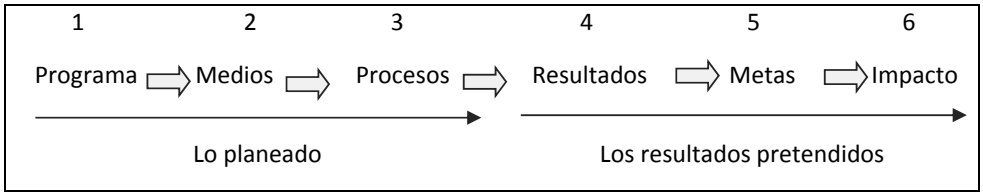

Figura 3. Diagrama de la secuencia de acontecimientos para conseguir los cambios.

Los resultados pretendidos mediante la aplicación del programa deben comprobarse mediante la comparación con la situación de partida y lo conseguido. Tras la realización de las actividades del programa se consiguen unos resultados inmediatos. Una vez finalizado el programa, los resultados son fruto de las interacciones de las distintas estrategias y actividades a lo largo de la aplicación del programa. Y aquellos que se mantienen en el medio y largo plazo, y que se incorporan como habilidades desarrolladas en las personas, utilizables en situaciones similares y distintas, que interactúan por asimilación y transferencia en otras situaciones, son conocidos como el impacto atribuible a la aplicación del programa. El impacto puede medirse en relación con la situación inicial o de partida, con los resultados inmediatos conseguidos por el programa y con los resultados mantenidos, extendidos, transferidos más allá del momento temporal en el que acaba el programa, y más allá

de la situación en la que pretendía incidir el programa, entendiendo que los cambios directos, como resultado de las actividades realizadas por el programa, a su vez contribuyen o generan otros cambios y comportamientos de segundo orden, pretendidos, pero también no pretendidos. Los resultados de segundo orden son el objeto de la evaluación de impacto.

4. LA EVALUACIÓN DEL PROGRAMA

Entendemos por programa un plan intencional y sistemático de acción, específica y expresamente desarrollado para servir a un grupo de metas y objetivos de calidad, tendentes a resolver una situación problemática o a mejorar una necesidad —basándose en la evaluación inicial de necesidades—, dirigido a atender a personas y grupos en un contexto determinado, debiendo estar fundamentado en teorías y modelos que han demostrado que ayudan a solucionar determinados problemas mediante los procedimientos que el programa aplica (Pérez Juste, 2006: 220).

El modelo lógico sugiere al evaluador estas tres ideas en relación con el programa:

1. Debe estar escrito en un lenguaje claro, preciso y concreto.
2. Debe contar con los recursos mínimos imprescindibles para su aplicación.
3. Debe centrarse en su aplicación, en las actividades prioritarias para conseguir los resultados clave que pretende conseguir.

El conocimiento del programa sirve de guía para explorar cómo se está aplicando para proceder a su evaluación. Y el modelo lógico sirve al propósito de analizar y valorar la coherencia entre la teoría que fundamenta el programa, para atender unas necesidades, y el despliegue empírico para conseguir los cambios que pretende (Greene, 2013: 72), a lo largo de un período de tiempo en un contexto dado.

La evaluación se configura como un principio activo, dinamizador de la actividad bien hecha, que promueve el análisis sobre la acción, la reflexión, la autocrítica.

El diseño del programa, para atender a unas necesidades previamente evaluadas, debe hacerse teniendo en cuenta el análisis de otros programas, para identificar las mejores prácticas y las investigaciones sobre las mismas, así como las mejores estrategias, y actividades, para conseguir los resultados pretendidos.

La presentación del programa, siguiendo el enfoque lógico, ofrece oportunidades para la descripción sistemática de lo pretendido por el programa, lo cual ayuda a su comprensión y facilita su aplicación. Posibilita centrarse en las metas y en las estrategias y actividades para su logro, y ayuda a detectar y corregir posibles lagunas y deficiencias. Muestra las relaciones entre los insumos del programa, las actividades y los resultados, haciendo explícita la teoría que subyace en el programa, a través de sus seis componentes: situaciones, insumos, actividades, resultados pretendidos, concepciones y factores externos. Y todo ello ayuda a desarrollar la comprensión del programa, mejorar su planificación y sus actividades, clarificar sus metas, centrar lo que debe ser evaluado y comunicarlo a los promotores y grupos de interés.

La planificación de la implementación del programa debe centrarse en su aplicación siguiendo el plan previo, y en la identificación de información clave y los procedimientos para su recogida, para la mejora del programa, de modo que ayude a priorizar los elementos críticos para orientar la mejora y ajustes del programa.

La evaluación continua del programa durante su aplicación, en su función de evaluación formativa, facilita la mejora del diseño del programa, la adecuación de los recursos necesarios y la planificación y realización de los procesos, así como la información acerca de las metas conseguidas tras su aplicación.

En la Tabla 2 se muestran las relaciones entre los elementos fundamentales de un programa y los criterios que deben garantizar su éxito (W. K. Kellogg F., 2004, p. 11).

Tabla 2. Relaciones entre el programa, los criterios y los beneficios del modelo lógico.

Elementos del programa	Criterios para garantizar el éxito del programa	Beneficios de usar el modelo lógico para la evaluación del programa
Diseño del programa, fundamentado, para atender las necesidades	Correcta definición de las metas y objetivos del programa, así como de sus posibles efectos a lo largo del tiempo.	Facilita comprobar las lagunas existentes entre los principios o lógica del programa, y ayuda a resolverlos. Clarifica las concepciones para que el éxito sea más probable. Promueve la comprensión y el consenso.
Planificación y recursos para su implementación	Las metas y objetivos del programa deben ser plausibles y posibles de conseguir.	Construye una comprensión compartida sobre el programa en su totalidad y de cómo debe funcionar cada una de sus partes. Y facilita la comprensión del uso efectivo de los recursos.
Implementación y gestión del programa	Debe permitir obtener información relevante, creíble y útil sobre la realización del programa.	Centra la atención en la gestión sobre las relaciones más importantes entre las acciones y los resultados.
Evaluación Formativa	La evaluación continua del programa durante su aplicación.	Facilita la mejora del diseño del programa, la adecuación de los recursos para la planificación y realización de los procesos.
Evaluación sumativa	La evaluación de la eficacia, de los resultados, del éxito del programa.	Ayuda a clarificar qué se debe evaluar y cuándo, de modo que los recursos de la evaluación sean utilizados adecuadamente. Información acerca de las metas conseguidas tras su aplicación, y la resolución de las necesidades o problemas.
Informes de los resultados	Identificación de los que van a usar los resultados de la evaluación.	Proporciona un modo de implicar y comprometer a los responsables en el diseño, procesos y uso de los resultados de la evaluación. Resume la complejidad de los programas para comunicar a los promotores, financiadores y demás grupos de interés.

La argumentación sobre los «qués», los «cómos» y los «porqués» del éxito de un programa deben basarse en evidencias y el informe de la evaluación debe reflejar de modo específico los pasos mediante los cuales se han conseguido los resultados. Informar, aclarar y compartir con las personas implicadas en el programa nuestra concepción de la evaluación, como evaluadores, las funciones que cumplirá y los procedimientos que vamos a utilizar para realizar la evaluación, incluidos los métodos de investigación y los instrumentos para la recogida de información y su posterior análisis, ayudará a evitar conflictos durante el proceso de la evaluación y contribuirá a la aceptación de los resultados y las recomendaciones de mejora de los programas y de la organización para su aplicación.

La evaluación del éxito del programa, de sus logros, o evaluación sumativa, descansa sobre el diseño claro del propio programa y de la concreción de las metas que pretende conseguir para la resolución de las necesidades o problemas a través de los procesos de su aplicación, además de en la recogida de información clave, mediante la evaluación continua formativa durante todas las fases anteriores, integrándolas con los resultados finales obtenidos por el programa en un informe descriptivo, comprensivo y valorativo.

5. LOS CRITERIOS DE CALIDAD

En el centro de los programas, de las acciones de intervención para conseguir cambios, deben situarse los valores que orientan las metas a conseguir. Y la evaluación deben regirse por los mismos criterios fundamentales, en relación con las metas y objetivos que persigue el programa, en coherencia con las necesidades que pretende resolver —mérito—, sus planes de acción, sus realizaciones y sus resultados y la calidad con la que el programa los consigue —valor—.

La afirmación de las metas proporciona cuestiones, preguntas e indicadores para realizar la evaluación de contexto, para recoger y aportar información para validar y mejorar las metas. La planificación proporciona cuestiones, preguntas e indicadores, sobre como conseguir las metas, generando cuestiones para realizar la evaluación de entrada

que proporcionará juicios sobre los planes que han sido elaborados para conseguir las metas, y cómo mejorarlos. Las acciones que debe realizar el programa para lograr esas metas iniciales aportarán cuestiones para realizar la evaluación de los procesos, las cuales proporcionarán información para la elaboración de juicios sobre las actividades para ayudar a mejorar las estrategias de las personas que aplican los programas. Los logros, la falta de logros y los efectos secundarios atraen la atención de las evaluaciones de producto, dirigen las preguntas a realizar sobre la evaluación de producto, que en última instancia permiten emitir juicios acerca de los resultados e identificar las necesidades para lograr mejores resultados en las sucesivas aplicaciones de los programas.

Las teorías en las que se fundamentan, ayudará a operativizar los objetivos en tareas y actividades para su aplicación con el fin de conseguir los resultados, que es la razón de ser del programa. Todo ello orientará la elaboración de indicadores representativos del programa, de su aplicación y de la evaluación de sus resultados en los sujetos, para la evaluación del programa.

Otros criterios a tener en cuenta en la evaluación de programas son el respeto por las exigencias legales, ya sean de la administraciones educativas, ya de los promotores y patrocinadores de los programas y de la sociedad, dado que hay programas que se crean por un mandato legal, con fondos públicos o con subvenciones privadas, con una finalidad determinada, para solucionar una necesidad y que están obligados a presentar un plan de evaluación y a realizar un informe de sus resultados. Tienen que rendir cuentas de la confianza que se les ha otorgado.

Los evaluadores deben conocer los criterios de calidad con los que debe trabajar el programa para conseguir sus objetivos, para valorar la adecuación a las necesidades que pretende resolver, su pertinencia, su mérito, como prueba resultante del proceso, y la calidad con la que el programa lo consigue, su valor, como conjunto de cualidades por las que el programa es apreciado, y el impacto que consigue a lo largo del tiempo, influenciados por los efectos del programa, que le conferirá un carácter específico. (Fernández Díaz, 2013: 121).

La esencia de la evaluación es el valor. Los valores deben orientar toda evaluación y son, por ejemplo, aquellos que afirman que todos los escolares deben conseguir una formación básica en conocimientos, habilidades, desarrollo personal, desarrollo de la capacidad para continuar el desarrollo personal y profesional a lo largo de la vida, el desarrollo del respeto a los derechos humanos, atención a las necesidades específicas de cada niño, el desarrollo de ciudadanos que aseguren la equidad de oportunidades efectivas para las personas.

En los modelos de evaluación de la calidad de los programas y de las organizaciones concurren los objetivos de evaluación para la mejora, para el desarrollo, y para el aseguramiento de la calidad al servicio de la mejora y del rendimiento de cuentas hacia la sociedad, basados en la confianza mutua y en la transparencia.

6. LAS METODOLOGÍAS DE INVESTIGACIÓN A UTILIZAR PARA LA EVALUACIÓN DE PROGRAMAS

Si la puesta en marcha de un programa se caracteriza por un proceso de adaptación a las condiciones reales, necesidades e intereses, los métodos utilizados para estudiar esa aplicación deben estar orientados a su descubrimiento, tratando de comprender los procesos desarrollados y los cambios producidos por el programa. La elección de los métodos en la investigación científica debe estar guiada por las cuestiones y propósitos que pretende resolver la investigación. Y es que la evaluación de programas, hecha con rigor, se convierte en investigación evaluativa, en una forma de investigación aplicada, empírica, realizada de manera sistemática y mediante una cuidadosa recogida de datos y análisis rigurosos, cuyo propósito es informar a la acción, apoyar la toma de decisiones y aplicar el conocimiento para resolver problemas humanos y sociales.

La investigación evaluativa, en coherencia con la complejidad de los programas a los que debe atender, desde el propio programa, en su aplicación a lo largo de su vigencia y atendiendo a las condiciones de «campo» en las que se aplica —centros educativos u otros—, que requiere contar con un amplio grupo de métodos y técnicas de investigación.

6.1. Las metodologías de investigación mixtas para atender la complejidad

Los diseños de la investigación evaluativa combinan metodologías cualitativas y cuantitativas; de ese modo los enfoques de metodología mixtos son más comprensivos por incluir metodologías cualitativas, como la observación participantes durante las sesiones en las que se está aplicando el programa, entrevistas semi-estructuradas, grupos de discusión representativos de los diferentes intereses y responsabilidades, junto con metodologías cuantitativas, con el uso de pruebas para la recogida de información previa y posterior al tratamiento o aplicación del programa, así como de cuestionarios para la valoración de la satisfacción y seguimiento de los resultados de los planes de mejora y análisis y valoración de sus resultados en el medio y largo plazo, o su impacto.

La utilización de diseños de grupo para la evaluación de los resultados de los programas y su impacto, con fines de inferencia causal generalizada, es uno de los muchos enfoques que han mostrado su utilidad para estudiar los problemas educativos y sociales, aplicando estas metodologías con todo el rigor que las caracteriza. Dado que los programas educativos son intervenciones de campo «socialmente complejas», no siempre fáciles de manipular, controlar y estudiar en las organizaciones educativas abiertas, lo que vulnera los supuestos teóricos de los diseños experimentales y cuasi-experimentales, es por lo que se necesitan diseños de investigación que tengan en cuenta esta complejidad, que aporten evidencias exhaustivas y utilizables para los interesados que buscan entender los resultados, mejorar sus programas y las condiciones sociales a las que atiende.

Los interesados en conocer cómo funcionan los programas realizan preguntas complejas, como la siguientes:

— ¿Qué es lo que funciona en el programa?

— ¿Cuándo y cómo funciona?

— ¿Podemos extender la aplicación de un programa y replicar sus resultados?

— ¿Cuáles son los costos y las consecuencias de los programas alternativos?

Preguntas difíciles de contestar con una única metodología, por lo que se necesitan diseños multi-método y multi-fase, para reunir evidencias significativas para satisfacer las necesidades de las partes interesadas en los contextos reales donde se desarrollan los programas. (Chatterji, 2005: 4).

La investigación evaluativa es directamente pertinente para comprobar el funcionamiento de los programas educativos o sociales en la acción, caracterizados por su complejidad y duración, así como por la complejidad de las situaciones del medio que los rodea, y se orienta a la toma de decisiones de cambio y mejora.

6.2. La elaboración de instrumentos para la recogida de información

Para la recogida de información deberemos elaborar indicadores representativos del programa, de las estrategias y actividades diseñadas para su aplicación, de su funcionamiento y de los resultados que pretende conseguir. Los indicadores de cada uno de los momentos en el desarrollo del programa deberán ser incluidos en los diferentes instrumentos de recogida de información para permitir recoger evidencias del programa, de su aplicación y de sus resultados.

La elaboración de instrumentos para la recogida de información en la evaluación de programas, supone conocer:

— Los principios psicopedagógicos del programa y las teorías de aprendizaje subyacentes.

— El plan para su aplicación en su contexto.

— Las metodologías de enseñanza y aprendizaje que utiliza y las actividades para su desarrollo del programa, y su comprobación.

— Los rendimientos conseguidos en relación con los objetivos del programa de estudio.

— Las características del contexto en el que se desarrolla.

— El clima social que generan sus miembros.

— La satisfacción de los usuarios —los profesores, los alumnos, los padres y la comunidad del entorno próximo—.

— El impacto que generan los resultados del programa en el medio y largo plazo y su alcance.

La información recogida mediante indicadores de comprobación del programa, basada en hechos y datos reales, una vez analizada e interpretada, será útil para informar a todos los que tienen derecho a saber acerca de cómo ha trabajado el programa, si se ha aplicado de acuerdo con el plan previsto —si se ha modificado explicar por qué—, y los resultados que ha conseguido en relación con las metas pretendidas —si ha conseguido otras y por qué—. Ello proporcionará información e indicaciones precisas sobre la conveniencia de la continuidad de su aplicación y la recomendación de su expansión, junto con los cambios y mejoras en aquéllos aspectos que lo necesite.

El programa ha debido ser implementado durante un tiempo suficiente para que haya podido lograr los efectos pretendidos y no se den dudas de que los elementos críticos en relación con los objetivos han sido realizados. De otro modo, si la aplicación del programa tiene poco que ver con lo que pretende conseguir, estaremos evaluando otra cosa distinta del programa. Por ello, además de la comprobación de la adecuación del programa en sí, para abordar los problemas que pretende resolver, es importante realizar la función de seguimiento de la aplicación del programa, o monitoreo, teniendo por objeto resolver dos interrogantes principales: Primero, si el programa alcanza o no a la población objeto adecuadamente; Segundo, si la realización de los procesos es congruente o no con las especificaciones de los objetivos y del diseño del programa. La monitorización en las evaluaciones es una ayuda necesaria para valorar su repercusión, pues los errores en el programa a menudo se deben a las ejecuciones incompletas o defectuosas de las intervenciones, más que a la ineficacia del propio programa, y aporta información en la decisión de continuar, ampliar o terminar con la aplicación de los programas.

Una síntesis sobre los momentos y dimensiones del desarrollo de los programas y las técnicas más adecuadas para la recogida de información para su evaluación, se muestran en la Tabla 3.

Tabla 3. La recogida de información en la evaluación de programas.

Componentes del programa	Dimensiones	Objeto	Técnicas e instrumentos
Análisis de necesidades	• El contexto	• Las personas y sus contextos	• Análisis de los resultados de evaluaciones previas • Lectura de informes, memorias • Realización de encuestas, entrevistas, observaciones.
El programa en su conceptualización y estructura	• Calidad de las metas, en relación con las necesidades	• Principios y teorías • Metas • Contenido • Estructura • Viabilidad	• Análisis de la conceptualización, estructura y contenido del programa. • Comparación con otros programas que han sido eficaces, en condiciones y problemas similares.
Planificación de la puesta en acción	• Contexto, • Recursos, • Estrategias para su aplicación	• Estrategias • Actividades • Clima • Criterios • Referencias	• Análisis de los recursos del contexto • Evaluación de la planificación • Evaluación de los recursos • Comparación con criterios de calidad para la implementación de programas
Análisis de los procesos	• Clima • Medida logros • Valoración	• Constatación • Criterios • Referencias	• Evaluación de los ritmos del programa, de la participación y satisfacción observación, entrevistas, cuestionarios
Análisis de los resultados y de su impacto	• Inmediatos • Satisfacción de las necesidades • Impacto a medio y largo plazo	• Continuidad • Cambio y mejoras • Terminación	• Comparación con criterios de calidad • Efectos del programa • Evaluación de los logros basado en los criterios • Evaluación de la satisfacción • Diseños cuasi-experimentales • Análisis coste-beneficio y efectividad
Propuestas de mejora	• Decisiones sobre los distintos elementos del programa	• Programa • Personas • Contextos • Recursos	• Informe y propuestas de mejoras • Planificación de su aplicación y seguimiento

Las evaluaciones deberán documentar plenamente sus propósitos negociados y los diseños y procedimientos aplicados y resultados. La propia evaluación debe ser evaluada interna y externamente mediante su comparación a través de normas para que su realización sea apropiadamente guiada y los interesados puedan examinar sus fortalezas y debilidades rigurosamente.

> El objetivo de la meta-evaluación es asegurar la calidad de los servicios evaluativos, evitar enfrentarse a prácticas ilegales o a servicios que no son de interés público, señalar el camino para el perfeccionamiento de la profesión y promover una mayor comprensión de la profesión evaluativa (Stufflebeam & Shinkfield, 1987: 52).

Identificar las necesidades de información de los responsables de los programas, documentar y analizar los contextos en los que se desarrollan, revelar la teoría subyacente del programa para iluminar las variables claves y sus relaciones, examinar la consistencia y autenticidad de los procesos de realización de los programas frente a la teoría del programa, proporcionar a los responsables información a tiempo para modelar y mejorar el programa en sus incipientes estadios, realizar estudios de resultados y de impacto y realizar estudios controlados de análisis de impacto, caracterizan la evaluación como investigación.

Junto con los resultados inmediatos conseguidos por el programa, debemos considerar los resultados en el medio y largo plazo, que se manifiestan a lo largo del tiempo, atribuibles a los efectos del programa, y que son conocidos con el nombre de impacto. El impacto de un programa tiene que ver con los resultados alcanzados por dicho programa, y estos dependerán, a su vez, de todo lo que le precede: las necesidades o problema a resolver, el programa diseñado para resolverlos, sus recursos, su aplicación y sus resultados.

La aplicación de las normas cumple con la función de meta-evaluación, cuyo significado se puede definir, por analogía con el término evaluación, como una investigación sistemática encaminada a emitir un juicio valorativo acerca de la calidad y los méritos de una evaluación. Atendiendo a la defensa de Scriven (1967), en relación con la meta-evaluación, ésta debería ser utilizada por todos los profesionales, con el fin de asegurar la calidad y honradez, precisión y veracidad del trabajo de evaluación realizado, y debe reflejarse en sus informes.

6.3. Compromiso con el rigor y la responsabilidad en la investigación evaluativa

Lo cierto es que, por definición y consenso en la profesión, la investigación evaluativa no sólo debe ser realizada con rigor y métodos de investigación adecuados, ya sean cualitativos o cuantitativos o ambos (normas de exactitud), sino que también con sentido de la responsabilidad y servicio hacia los futuros usuarios de la información (normas de utilidad y de propiedad) y de responsabilidad ante la sociedad (normas de responsabilidad), (JCSEE, 2011).

7. LA ELABORACIÓN DE LOS INFORMES DE LA EVALUACIÓN DE PROGRAMAS

La finalidad de los informes es comunicar los resultados de la evaluación del programa, y los informes de las evaluaciones deben documentar plenamente sus propósitos negociados y los diseños y procedimientos aplicados y sus resultados.

Tras el proceso de comparación entre los criterios y el funcionamiento real del programa, se deben integrar los resultados en un juicio valorativo final, de modo que sirva para orientar en la mejora del programa y de las personas implicadas en él, en aquéllos componentes y elementos que más se necesite. El informe debe integrar la valoración de la totalidad del programa.

La información evaluativa debe aportar conocimiento, enseñanza, sobre los problemas sociales y sus concepciones y sobre la teoría implícita de los programas. Y debe atender las diferentes necesidades de información de los diversos grupos implicados en el programa, se deben utilizar diversos estilos de comunicación. Por eso, los informes de evaluación deben mostrar:

— Las cuestiones que han aconsejado la evaluación.

— Los procesos seguidos para realizar la evaluación.

— Las fuentes de donde toman sus datos en los diferentes momentos de la evaluación.

— Los métodos, técnicas e instrumentos utilizados, de acuerdo con el objetivo y la dimensión a evaluar.

— Los criterios con los que se comparan los programas, su funcionamiento y sus resultados.

— Los criterios en los que se basa que llevan a emitir los juicios valorativos sobre el programa, su implementación y sus logros.

— Las recomendaciones que se derivan de la evaluación.

— El plan para incorporar las recomendaciones al programa y las indicaciones para el seguimiento de su puesta en práctica.

El informe para la comunicación de los resultados del programa debe facilitar la comprensión del funcionamiento del mismo, siendo esto clave para el aprovechamiento de sus resultados para la mejora de los programas y para lograr apoyos de las autoridades para su continuidad, en su función de evaluación como rendimiento de cuentas. Es importante que se realice la diseminación de los resultados para ayudar a otros grupos o poblaciones y contextos con similares necesidades y problemática.

8. UTILIZACIÓN DE LOS RESULTADOS DE LA EVALUACIÓN

La evaluación, como parte de la investigación social, debe conocer y estar al servicio de atender las necesidades de la sociedad, que guiaran las cuestiones de investigación. Los evaluadores, como investigadores, deben conseguir que los responsables sociales y políticos se impliquen para identificar y orientar las cuestiones que guiarán la investigación y conseguir su implicación en los procesos de investigación como una forma de asegurar la utilidad, la aplicación y la difusión de los hallazgos de la evaluación. La implicación de los políticos y la comunidad objeto del estudio evaluativo, garantizarán la utilidad de los hallazgos de la investigación, para la comunidad concreta en la que se realiza y en otras similares.

La evaluación no es sólo una actividad técnica sino que es, también, una investigación diseñada para dirigir el proceso de formación de la

comunidad en la que se desarrolla. Por ello, el evaluador debe ser consciente de que los resultados de la evaluación deben ser comunicados a todos los que tienen derecho a conocerlos y debe llegar allí dónde se toman las decisiones, de un modo adecuado y a tiempo.

Los resultados de las evaluaciones tienen como finalidad contribuir a la toma de decisiones de cambio y mejora de los programas, de la organización en la que se desarrolla, del personal que los aplica y de los que se benefician de ellos. Pero sólo repercutirá en el proceso de toma de decisiones si resulta una información convincente, coherente, adecuada al problema que pretende resolver y al contexto y oportuna, debiendo tener en cuenta los recursos necesarios para poder llevar a cabo las reformas del programa y los cambios organizativos y formativos que de los resultados de la evaluación se deriven.

La información evaluativa debe aportar nuevas luces sobre las concepciones y las teorías implícitas de los programas y de sus actividades. Debe informar a los que han encargado la evaluación y a todos los que intervienen en el programa sobre cómo mejorar el programa, en función de los cometidos para los que se creó y las necesidades actuales y futuras detectadas.

Los resultados de las evaluaciones sistemáticas, hechas con rigor y fundamentadas en la teoría y en la metodología, tienen repercusiones de tipo conceptual, persuasiva e instrumental:

— La repercusión conceptual hace referencia al poder de las evaluaciones para influir en el pensamiento de los que deben utilizar sus conclusiones, de modo que contribuya a la mejora del desarrollo de los programas y de la política y estrategias para implementarlos, debiendo repercutir de una manera directa y positiva en sus beneficiarios, tanto internos como externos y con impacto importante sobre la comunidad.

— La utilidad persuasiva de los hallazgos de la evaluación tiene que ver con los argumentos utilizados en la exposición de los resultados, con un claro esfuerzo por defender su utilidad y las ventajas que ello supone para la comunidad.

— La utilidad instrumental de los resultados de una evaluación tiene que ver con cómo el evaluador debe presentar el informe de los resultados acorde con los estilos cognitivos de los que toman las decisiones y con cómo entiende y respeta las obligaciones y compromisos de los patrocinadores y promotores de los programas, arbitrando su participación en el proceso del diseño de la evaluación y recogiendo sus intereses y sensibilidades. Tiene que ver con la presentación de los resultados de la evaluación a tiempo, cuando se necesitan. Por ejemplo, en un centro educativo, los resultados de la evaluación deberían estar disponibles antes de finalizar el curso, para poder informar al consejo escolar y poder tomar decisiones antes del comienzo del nuevo curso, tanto por parte de la organización del centro como por parte de los equipos docentes, familias y alumnos. Asimismo, las recomendaciones para la utilización y la diseminación de los resultados de las evaluaciones deben formar parte de los diseños de la evaluación, de modo que facilite su aplicación a los usuarios de los programas.

La utilización de los resultados de la evaluación requieren de su difusión, siendo ello responsabilidad de los investigadores evaluativos. A quienes encargan la evaluación se les debe proporcionar el informe de los resultados, junto con las propuestas de cambio y mejora, teniendo en cuenta las personas que han de llevarlas a cabo y los recursos formativos y de otro tipo que se requieren para que las propuestas tengan éxito.

El evaluador también debe ser un diseminador de la información de varios modos: mediante la publicación de informes técnicos especiales que sirvan a otros evaluadores y a través de su presentación en foros de interés, y mediante la utilización de su experiencia en sucesivas evaluaciones.

9. CONCLUSIONES

El potencial de la evaluación para la mejora de la calidad de los programas, de las organizaciones, de las personas y de la sociedad es de una gran trascendencia al ayudar a identificar los aspectos positivos y negativos con los que operan las organizaciones y los diferentes elementos que

contribuyen al logro de sus resultados. La idea central de las evaluaciones es proporcionar información fundamentada que pueda ayudar a los responsables de los programas a recoger información regularmente para mejorar los programas, hacer uso efectivo y eficiente de los recursos, a tiempo, con el fin de servir las necesidades especificadas en el programa, e informar a todos los beneficiarios adecuadamente y con equidad. (Sttuflebeam & Shinkfield, 2007: 330).

Los vocablos de «calidad» y «evaluación» adquieren una fuerte carga simbólica, utilizados en relación con lo social y profesionalmente correcto, representando lo mejor, lo deseable lo correcto. Ambos conceptos se presentan conexionados, y no es extraño, ya que la calidad exige evaluación y la evaluación se justifica como garantía de calidad (De la Orden, 2009: 18). Garantizar o asegurar la calidad y la necesidad de rendir cuentas, ha potenciado el uso de los conceptos y funciones de la planificación y de la evaluación de programas, de sus metodologías y sus procedimientos.

Desde la perspectiva de la evaluación, un modelo lógico atiende al desarrollo de los componentes de la evaluación y a la forma en que se armonizan con los objetivos generales de los programas y de sus teorías subyacentes, ayudando a priorizar y enfocar las cuestiones de interés. La evaluación debe estar guiada por la teoría del programa.

La evaluación de programas, aplicada en todo su alcance, constituye una estrategia de aprendizaje continuo para la mejora de los programas, de las personas, de la propia institución u organización en la que se desarrollan y de la sociedad a la que sirve.

EJERCICIOS DE AUTOEVALUACIÓN

Seleccione un programa que conozca y esté interesado en conocer cómo se aplica y los resultados que consigue, y realice la evaluación del mismo, guiándose por las siguientes cuestiones:

1. Elabore un diagrama en el que describa las necesidades o razones del programa, problemas que pretende resolver, sus metas y objetivos, los resultados que afirma pretender conseguir el programa, los factores

personales, de recursos y de contexto necesarios y aconsejables para que el programa pueda conseguir sus objetivos, las estrategias y actividades que debe realizar el programa en su aplicación y los resultados a conseguir, orientados al logro de las metas finales del programa y su impacto.

2. Describa, de modo breve pero suficiente, fundamentado, cada uno de los elementos mencionados en el diagrama anterior, y cómo debe funcionar cada elemento para conseguir las metas y objetivos pretendidos.

3. Determine los criterios de calidad del programa, de su aplicación y de sus resultados inmediatos y en el medio y largo plazo, en función de lo que el programa pretende conseguir y el grado de calidad con que debe conseguirlo. Concrete los niveles mínimos que debe lograr el programa para considera que funciona bien. Expréselo mediante una tabla de correspondencias que muestre lo que va a ser objeto de evaluación, y los criterios para garantizar su éxito y calidad.

4. Indique, para cada elemento del programa, la metodología y técnicas de recogida de información a utilizar para su evaluación, comprobación de su funcionamiento y resultados. Muéstrelo mediante una tabla.

5. Continuando con el punto anterior, referido a la metodología de investigación, explique los análisis de datos que aplicaría, y arguméntelo.

6. Elabore una tabla para aplicar las normas del «Joint Committee» a la evaluación del programas. Utilice un sistema de valoración tipo Likert, de 1, mínimo, a 5, máximo cumplimiento de la norma.

7. Realice la evaluación del programa siguiendo las estrategias elaboradas en los apartados 1 a 6.

8. Elabore el informe de la evaluación del programa, en el que se refleje todo el proceso seguido para evaluar el programa, sus resultados, conclusiones y recomendaciones para utilizar los resultados de la evaluación e introducir los cambios y mejoras necesarios en el programa.

9. Realice una valoración crítica de la práctica evaluativa realizada.

REFERENCIAS BIBLIOGRÁFICAS

Chatterji, M. (2005). «Evidence on "what works"»*:* An argument for extended-term mixed-method (ETMM) evaluation designs. *Educational Researcher,* 34,5, 14-24.

De La Orden, A. (2009). «Evaluación y calidad: Análisis de un modelo». *Estudios sobre Educación* (ESE), 16, 17-36. Disponible en:

http://dadun.unav.edu/bitstream/10171/9157/1/16%20Estudios%20Ea.pdf

Fernández Díaz, M. J. (2013). «Evaluación del impacto para un cambio sostenible en las organizaciones educativas». *Revista Española de Pedagogía*. 254, 19-138. Disponible en: <https://revistadepedagogia.org/lxxi/no-254/evaluacion-del-impacto-para-un-cambio-sostenible-en-las-organizaciones-educativas/101400010300/> Acceso en: 6 mar.2015.

Freedman, A. M., Simmons, S., Lloyd, L. M., Redd, T. R., Alperin, M., Salek, S. S., Swier, L.& Miner K. R. (2014). Public Health Training Center Evaluation: A Framework for Using Logic Models to Improve Practice and Educate the Public Health Workforce. *Health Promot Pract*. 15, 80-88. En: <http://hpp.sagepub.com/content/15/1_suppl/80S> Acceso: 8 oct. 2015.

Greene, J. C. (2013). «Logic and evaluation theory». *Evaluation and Program Planning,* v. 38, p. 71-73. doi:10.1016/j.evalprogplan.2012.03.017

Joint Committee on Standards for Educational Evaluation. (2010) *Program Evaluation Standards Statements*. En: <http://www.jcsee.org/program-evaluation-standards-statements> Acceso5/2/ 2014.

Mclaughlin, J. A. & Jordan, G. B. (1999). «Logic models: a tool for telling your program's performance story». *Evaluation and Program planning.* 22 (65-72). En: <https://www.researchgate.net/publication/222500540_Logic_Models_A_Tool_for_Telling_Your_Program»s_Performance_Story>. Acceso en: 4 jun. 2007.

Newton, X. A., Poon, R. C., Nunes, N. L., & Stone, E. M. (2013). «Research on teacher education programs: Logic model approach». *Evaluation and Program Planning*. 36, (88-96). En: <http://www.ncbi.nlm.nih.gov/pubmed/22990038>. Acceso en: 20 may. 2012

Pérez Juste, R. (2006). *Evaluación de programas educativos*. Madrid: La Muralla.

Scriven, M. (1967). «The methodology of evaluation». In R. Tyler, R. Gagne and M. Scriven. Perspectives on Curriculum Evaluation (*AERA Monograph Series – Curriculum Evaluation)* (Chicago, Rand McNally and Co). En:

http://files.eric.ed.gov/fulltext/ED030949.pdf

Scriven, M. (2015). *The logic of evaluation. Ontario Society fo the Study of Argumentation.* OSSA Conference Archive. Paper 138,16, 2007. En: http://scholar.uwindsor.ca/ossaarchive/OSSA7/papersandcommentaries/138

Stufflebeam, D. L. & Shinkfield, A. J. (1987*). Evaluación sistemática. Guía teórica y práctica.* Madrid: Paidós, MEC.

— (2007). «Evaluation. Theory, Models and Applications». San Francisco, CA: John Wiley and Sons. Inc.

W. K. Kellogg Foundation (2004). *Using logic models to bring together planning, evaluation, and action: Logic model development guide*. Battle Creek, MI. En:

https://www.wkkf.org/resource-directory/resource/2006/02/wk-kellogg-foundation-logic-model-development-guide

Tema 6

Las Normas de calidad del Comité Conjunto para la evaluación de programas educativos

1. Objetivos y resultados de aprendizaje.
2. Introducción.
3. Características y conceptos de los estándares de calidad de la evaluación de programas.
   - 3.1. Características de las Normas en su tercera edición.
   - 3.2. Los responsables de los programas en los centros educativos.
   - 3.3. Conceptos y definiciones sobre los estándares para la evaluación de programas.
4. Las Normas para la evaluación de programas del Joint Committee.
   - 4.1. Normas de utilidad.
   - 4.2. Normas de viabilidad o factibilidad.
   - 4.3. Normas de probidad, o adecuación legal y ética.
   - 4.4. Normas de precisión o exactitud.
   - 4.5. Normas de responsabilidad.
5. Utilización de las Normas del Joint Committee para el desarrollo de los programas, de su evaluación y de su responsabilidad pública.
6. Lista de control de las Normas para la comprobación de su aplicación en la evaluación de los programas.

Ejercicios de autoevaluación.

Referencias bibliográficas.

1. OBJETIVOS Y RESULTADOS DE APRENDIZAJE

Pretendemos que el alumno, con el estudio de este tema, comprenda y valore la finalidad de las normas para la evaluación de programas y su utilidad para la realización de evaluaciones como futuro evaluador o usuario de evaluaciones de programas.

De modo concreto, se espera que el alumno, al terminar el estudio del tema, consiga los siguientes objetivos de aprendizaje:

— Conocer y valorar la finalidad de los estándares de calidad del Joint Committee para la evaluación de programas.

— Conocer y valorar las cinco características o atributos de calidad de los estándares de la tercera edición del 2011.

— Valorar la aplicación de los estándares de calidad y sus normas correspondientes para la realización de evaluaciones de programas.

— Aplicar los estándares de calidad del Joint Committee en la evaluación de programas.

— Valorar y argumentar la utilidad de los estándares para la mejora de las evaluaciones de los programas, de los propios programas y de la educación.

2. INTRODUCCIÓN

Los estándares de calidad para la evaluación de programas educativos forman parte de un proyecto dirigido por Stufflebeam como coordinador del Joint Committee para el desarrollo de estándares para la eva-

luación de programas que se inició en 1975. Representa una de las asociaciones más importante de profesionales concernidos con la calidad de la evaluación educativa en los Estados Unidos de América y en Canadá. Su objetivo es guiar las evaluaciones de programas para asegurar que éstas sean útiles, factibles, éticas y precisas, y desaconsejar su evaluación si no cumplían con estos cuatro criterios de calidad de las evaluaciones. En la actualidad se ha añadido un quinto criterio relacionado con la responsabilidad de la calidad de las evaluaciones.

Los autores de las normas aconsejan que los evaluadores, como profesionales de la evaluación, y los beneficiarios de las evaluaciones cooperen entre sí para que las evaluaciones puedan realizarse, aplicando las normas del Joint Committee, con la certeza de que su realización contribuirá a la mejora de la educación.

La finalidad de las normas o estándares para la evaluación de programas, proyectos y materiales del Joint Committee es que sirvan de ayuda a los prácticos de la evaluación de programas educativos, siempre con la colaboración de los expertos en evaluación de programas educativos. Además de su objetivo para la evaluación, los estándares guían el diseño y aplicación de los programas, formando parte de un proceso continuo de diseño, aplicación, evaluación y mejora de los programas y mejora de la evaluación.

En la actualidad, el Joint Committee ha publicado normas para la Evaluación del Personal (segunda edición), Normas para la Evaluación de programas (tercera edición) y Normas para la Evaluación del Aula para profesores de Educación Infantil. El Joint Committee está acreditado por el Instituto Nacional Americano de Estandarización (ANSI), siendo los estándares del Joint Committee los únicos certificados por ANSI en América.

3. CARACTERÍSTICAS Y CONCEPTOS DE LOS ESTÁNDARES PARA LA EVALUACIÓN DE PROGRAMAS

La primera publicación del *Joint Committee on Standards for Evaluations of Educational Programs, Projects, and Materials*, *Standards for Educational Evaluation*' (JCSEE) en los EEUU, fue realizada bajo la

dirección de Daniel Stufflebeam. Ha habido una segunda edición de las normas, revisadas bajo la dirección de Sanders, en 1994, con traducción al castellano. Y una tercera revisión publicada en el 2011 coordinada por Yarbrough, Shulha, Hopson y Caruthers. El Comité se reúne periódicamente para revisar la aplicación de los estándares y actualizarlos de acuerdo con las necesidades detectadas, acordando no superar los 10 años para publicar nuevas ediciones con los cambios aconsejados.

El término inglés «*standard*» se traduce al español como estándar o norma. El documento del «*Joint Committee*» clarifica que las evaluaciones pueden ser realizadas sobre programas ya establecidos en funcionamiento, instituciones o áreas de servicios, o sobre proyectos. Con posterioridad a su primera edición, se amplió su aplicación a la evaluación de alumnos y profesores.

Los nombres de los estándares y sus definiciones, tienen el copyright de la JCSEE y están aprobados como una Norma Nacional Americana. El permiso se da libremente para su utilización con fines educativos y académicos con atribución a la JCSEE. Pueden consultarse las revisiones y aplicaciones actualizadas en la página web de «*The Evaluation Center*», *Joint Committee on Standards for Educational Evaluation, Joint Committee on Standards for Educational Evaluation* o en la dirección Postal en la Western Michigan University, Kalamazoo, Michigan 49008-5178.

3.1. Características de las Normas en su tercera edición

Los estándares para la evaluación de programas, en su tercera edición, proporcionan una guía integrada para la evaluación de programas que potencia el aprendizaje humano y su práctica a lo largo de la vida. Estos estándares se aplican a una amplia variedad de lugares y situaciones en los que el aprendizaje tiene lugar: escuelas, universidades, comunidades, organizaciones gubernamentales, organizaciones médicas y sanitarias, ejército, empresas del sector privado u organizaciones no gubernamentales, entre otras.

La revisión del 2011 mantiene el número de 30 normas, pero introduce un nuevo grupo sobre los cuatro anteriores, de modo que están organizadas en cinco grupos correspondiendo a cinco atributos claves sobre

la calidad de la evaluación. Cada grupo de normas, o atributo, es introducido por los conceptos claves que lo fundamentan, y para cada norma se proporciona una breve definición, clarificación, fundamento y sugerencia para su aplicación.

Los estándares se dirigen a servir como:

a) Guías para el diseño y realización fundamentados de evaluaciones y promover el uso de los resultados de la evaluación de modo apropiado.

b) Recursos para el aprendizaje de los promotores, concernidos e interesados en los programas y en su evaluación, acerca de los propósitos para la realización de evaluaciones y qué pueden esperar tras los esfuerzos realizados en la evaluación.

c) Marcos de trabajo para realizar metaevaluaciones o valoraciones sobre la calidad de las prácticas evaluativas en determinados proyectos y programas.

d) Recursos para el desarrollo de propuestas para la evaluación de programas en desarrollo o nuevos programas y proyectos.

e) Guía de criterios para la evaluación de los conocimientos del evaluador y de su credibilidad.

En la fundamentación de las normas en su tercera edición, se proporcionan orientaciones para fomentar una práctica reflexiva relacionado con:

— Evaluar, sí y cuándo.

— Cómo seleccionar a los evaluadores y otros expertos.

— El impacto de las culturas, los contextos y las políticas en los programas y en su evaluación.

— La comunicación e implicación de los patrocinadores y beneficiarios e interesados.

— Elementos técnicos para la planificación, diseño y gestión de las evaluaciones,

— Uso y mal uso de las evaluaciones.

— Otros temas relacionados con la calidad de la evaluación, la mejora y la responsabilidad —*accountability*. (Yarbrough, Shulha, Hopson & Caruthers, 2011: xii)—.

Como en las anteriores ediciones, la tercera edición revisada de las normas se basa en la transparencia y en un fuerte compromiso de los responsables de los programas y los beneficiarios de los mismos con la calidad. La tercera edición de las normas introduce importantes cambios. Por una parte, las normas de utilidad, viabilidad, propiedad o adecuación y exactitud mantienen su esencial importancia.

Con la finalidad de atender a las sugerencias de los usuarios de las normas y crear una mayor integración de las situaciones reales en las que se desarrollan los programas, la tercera edición de las normas proporciona un escenario para cada atributo o grupo de normas, con aplicaciones que se extienden a todas las normas relacionadas para ese escenario. Además, integra temas significativos tales como el papel de la cultura y el contexto en la evaluación, relacionando los atributos y sus normas con las situaciones concretas. (Ibíd.: xiii).

Un tema recurrente para la realización de evaluaciones de programas de alta calidad es la implicación de los interesados, no solo para conseguir que las evaluaciones de los programas sean de mayor utilidad y tengan un mayor impacto sobre la viabilidad y la exactitud o precisión de las evaluaciones, si no, también, por su papel esencial en la evaluación de la responsabilidad —*accountability*.

Asimismo, se incide de modo especial, en la importancia de una elevada calidad de la comunicación, comenzando por la necesidad de comunicar el plan de evaluación, mencionado repetidamente a lo largo de los atributos o categorías y sus normas. La cuestión no es solo que los informes formales e informales deben ser precisos, exactos y comunicados oportunamente, a tiempo, para que la evaluación sea útil, incluida la comunicación durante el proceso de la evaluación, en su función formativa; la comunicación de alta calidad es necesaria para afrontar los conflictos de intereses, el respeto a los derechos humanos, la selección y recogida de datos y la planificación y realización de las evaluaciones de calidad. De ese modo, se da mayor consistencia y coherencia entre cada

atributo y las normas que los desarrollan, atendiendo a los compromisos necesarios con los grupos interesados.

Un tercer cambio significativo de esta tercera edición revisada de las normas es el énfasis que se muestra en la necesidad de mejorar y asegurar la responsabilidad de las evaluaciones mediante la metavaluación sistemática. En esta nueva edición se incorpora un nuevo atributo, la responsabilidad en la evaluación. Es el quinto atributo, apoyado por tres normas, mediante las cuales se exhorta a una rigurosa documentación de las evaluaciones y a la realización de su metaevaluación interna y externa, con propósitos formativos y sumativos. Las normas señalan explícitamente que las evaluaciones sean metaevaluadas para la mejora de la evaluación y con propósitos de responsabilidad o rendición de cuentas.

El nombre de las normas incluidas en cada uno de los atributos que componen las totalidad de las normas para realizar evaluaciones de calidad, también han sido revisados con la finalidad de presentarlas de un modo más claro y conciso.

Los responsables de esta tercera edición invitan a utilizar las normas y a hacer sugerencias sobre la necesidad de cambios y mejoras, así como la conveniencia de seguir revisando y actualizando las normas en períodos de cinco años no debiendo exceder de los diez años.

3.2. Los responsables de los programas en los centros educativos

Los grupos implicados en los programas y en sus evaluaciones juegan importantes funciones en la evaluación, siendo algunos de estos roles los siguientes:

— Propuesta del grupo de revisores, evaluadores y programas a evaluar en una organización, dependiendo de sus necesidades de evaluación para comprobar su calidad en el programa, en su desarrollo y en sus resultados e impacto.

— Búsqueda de financiación para la evaluación de programas.

— Propuesta de audiencias que deben estar informadas de la evaluación del programa y sus resultados.

— Propuesta de los evaluadores, con experiencia profesional, para la evaluación sistemática de los programas.

— Propuesta de responsables en el diseño de programas y en la evaluación de programas para trabajar juntos en elaborar el plan de evaluación y concretar sus propósitos, metas, objetivos, incluyendo las características técnicas y dominio de los especialistas, y seleccionar el grupo de responsables de la institución que deben participar, representativos de los diferentes grupos de implicados concernidos en el programa y en su evaluación.

— Propuesta de metaevaluadores, es decir, aquellos que evalúan las evaluaciones del programa utilizando estos u otros estándares.

3.3. Conceptos y definiciones sobre los estándares para la evaluación de programas

En la tercera edición, como en las anteriores, los responsables de la coordinación de la revisión de las normas en el «*Joint Committee*», incluyen en su publicación algunas definiciones de interés que se utilizan en el texto, y que recogemos a continuación:

— ***Estándar*:** «Un principio acordado por gente comprometida en una práctica profesional, que, aunque la conocen, mejorarán la calidad y claridad de esa práctica profesional, por ejemplo, la evaluación» de programas, o el diseño de programas, o la aplicación de los programas. (JCSEE,1994: 5-6).

En la edición del 2011 de las Normas para la Evaluación de Programas, la palabra «estándar» es utilizada con dos características claves. La primera utiliza la palabra estándar en el sentido de identificar y definir una evaluación de calidad y guiar a los evaluadores y usuarios de la evaluación en conseguir realizar una evaluación de calidad. El segundo sentido de estándar indica que éstos no son leyes, sino afirmaciones desarrolladas mediante el consenso, considerando las peticiones y argumentaciones de los concernidos en los programas y sus evaluaciones, revisadas y

aprobadas por los miembros del JCSEE, siguiendo los procedimientos de la *American National Standards Institute* (ANSI).

Estas normas son similares a otros tipos de normas, como las normas técnicas de la *International Standardization Organización* (ISO), que especifican las dimensiones específicas exactas para cuestiones de la industria, las Normas de los Sistemas de Gestión de la Calidad de las Organizaciones, o Normas relativas a los procesos de aprendizaje y a sus resultados, identificando los estándares como modos para mejorar su calidad.

Las normas no especifican los procedimientos exactos a seguir en las diversas situaciones en las que pueden aplicarse, debido a que requiere una aplicación equilibrada de cada norma individual dependiendo de loas elecciones individuales atendiendo a las situaciones específicas. En las situaciones especificas de aplicación de los estándares para la evaluación de programas, las personas implicadas en la evaluación deben decidir como realizar la mejor evaluación basada en la prioridad de necesidades. Los estándares pueden ser aplicados en todas las evaluaciones, pero el modelo exacto sobre como se aplican puede diferir. (Yarbrough, Shulha, Hopson & Caruthers, 2011: xxiv).

— ***Programa Educativo y Proyecto Educativo***: En la segunda edición de las normas, en 1994, programa educativo es definido como «Actividades educativas que se dan de una manera continuada». Los ejemplos incluyen un programa de desarrollo de habilidades sociales para los alumnos de 4.º de primaria, un programa de educación vial para los alumnos de 2.º de primaria y un programa de estrategias de aprendizaje para los alumnos de 1.º de secundaria, entre otros.

— ***Proyecto educativo***: En la segunda edición de las normas, proyecto educativo es definido como «Actividades educativas que se dan en un período concreto de tiempo. Como ejemplos pueden ser un proyecto de formación sobre estrategias de enseñanza, de tres días de duración, al comienzo de un curso escolar, dirigido a profesores, un proyecto de desarrollo de potencialidades humanas para el trabajo colaborativo de dos meses de duración para desarrollar con los alumnos de segundo de carrera. Cuando un proyecto se institucionaliza, para su aplicación en sucesivos años, se denomina programa».

En la edición del 2011, tras la discusión entre gestores y administradores de programas y proyectos, se considera que no hay una distinción fundamental entre ambos conceptos en relación con su duración, dado que algunos proyectos pueden tener una duración mayor que algunos programas. Tampoco existen diferencias entre ambos conceptos en cuanto a diferencias de tipo cualitativo o de estructura y funciones. Normalmente los proyectos pueden proporcionar un marco de trabajo para los programas, y no al revés. De modo que en esta edición no hay una diferencia para aplicar los estándares a un proyecto o a un programa, únicamente se debe considerar la situación, el contexto, en el que se desarrollan.

Los programas son mucho más que las actividades que los desarrollan. Contienen múltiples componentes, siendo cada componente objeto de evaluación en sí mismo. Los componentes específicos que pueden ser evaluados son:

— Los contextos y como interaccionan con los programas y los componente de los mismos.

— Los participantes y otros beneficiarios, así como aquello que enfrenta costes o pérdidas de beneficios.

— Las necesidades, problemas y espacios de políticas de los programas y sus contextos.

— Las metas y objetivos.

— Los recursos y costes de todo tipo, incluyendo los de personal, instalaciones, materiales y costos de oportunidad, no presupuestados.

— Las actividades, procedimientos, planes, políticas y productos o resultados.

— Los modelos lógicos, las creencias, concepciones y teorías del programa implícitas y explícitas explicando por qué y cómo deben trabajar los programas.

— Los rendimientos, resultados, beneficios, logros e impactos.

La descripción completa de los programas incluye los elementos claves presentados arriba. Describimos los programas como la aplicación sistemática de los recursos guiados mediante los procesos lógicos, las creencias, y concepciones que identifican las necesidades humanes y los factores relacionados con ellas. (Ibíd.: xxiv).

Definimos de un modo completo «Programa» como:

— un grupo de actividades planteadas,

— que gestiona recursos,

— para conseguir las metas especificadas,

— relacionadas con las necesidades específicas

— de grupos o de individuos,

— en contextos específicos,

— resultando en logros e impactos documentados,

— siguiendo un sistema de creencias asumido, explícito e implícito, de acuerdo con teorías de intervención sobre como deben trabajar los programas,

— con específicos e investigables costes y beneficios.

Estos estándares del JCSEE no han sido diseñados para ser aplicados de igual modo a todos los programas, sino que más bien deben ser adaptados para servicios educativos, desarrollo de recursos humanos, salud, bienestar y otros tipos de programas en los cuales las metas incluyen cambios en la motivación humana, en las actitudes, conocimientos, destrezas y realizaciones. (Ibíd.: xxii-xxiv).

— ***Audiencia:*** «Individuos o grupos que están concernidos con la evaluación y pueden estar directamente implicados, por ser los responsables de los programas, o quienes los aplican, o estar afectados por una evaluación de programas».

— ***Evaluación:*** en anteriores ediciones, el concepto de Evaluación se define como «La investigación sistemática del valor o mérito de un objeto».

En la tercera edición se amplia la definición de ***Evaluación de Programas***, que incluye:

— La investigación sistemática de la calidad de los programas, proyectos, subprogramas, subproyectos, y cualquiera de sus componentes o elementos, juntos o de modo parcial.

— Para propósitos de toma de decisiones, juicios, conclusiones, hallazgos, nuevos conocimientos, desarrollo organizacional y capacidad para construir respuestas a las necesidades identificadas por las personas concernidas en los programas.

— Dirigido a la mejora y al rendimiento de cuentas de los programas y de los sistemas de los usuarios.

— Contribuyendo al valor organizacional y social del programa educativo. (Ibíd.: xxv).

El primer componente de la definición, investigación sistemática, hay que decir que en las ciencias sociales, la investigación se centra fundamentalmente en la necesidad de generar conocimiento creíble, generalizable. En una evaluación sistemática sobre un programa educativo, o social, o de sanidad, o medio-ambiental, o cualquier otro, las cuestiones de investigación se centran en la calidad del programa en concreto. Las audiencias principales para este tipo de investigaciones son los propios grupos de investigadores, y las audiencias secundarías serían aquellos que pudieran utilizar los resultados de conocimiento para su aplicación en la práctica.

La evaluación de programas utiliza todas las metodologías de la investigación de las ciencias sociales al servicio de la mejora de los programas y del cuerpo social en el que los programas se desarrollan. Las evaluaciones de programas ayudan a aquéllos que necesitan resolver cuestiones, dudas sobre cómo trabajan determinados programas y sus diferentes elementos o componentes. Los evaluadores utilizan sus habilidades para investigar tales cuestiones como el desarrollo de un programa, sus procesos, su teoría, su viabilidad, sus resultados y su impacto. Utilizan todas las metodología de investigación al servicio de las necesidades de evaluación de los programas y para aportar valor específico a los grupos interesados en dichos programas. (Ibíd.: xxv).

La evaluación de los programas y proyectos debe documentarse y evaluarse de modo que puedan mejorar y servir a la transparencia y responsabilidad. La evaluación acerca de la calidad de las evaluaciones es nombrada como meta-evaluación. Meta-evaluación no debe confundirse con meta-análisis. El meta análisis se refiere a la práctica de revisión de múltiples investigaciones realizadas sobre un mismo tema con el fin de obtener conclusiones generalizables basadas en la calidad de dichas investigaciones.

4. LAS NORMAS PARA LA EVALUACIÓN DE PROGRAMAS DEL JOINT COMMITTEE

El total de las normas o estándares es 30, y se agrupan en torno a cinco categorías relacionadas con la Utilidad de la información que aporte la evaluación; la Viabilidad, referida a que pueda realizarse la evaluación sin causar más problemas que beneficios; la Honradez o Adecuación, para que las evaluaciones sean conducidas de manera legal y ética; la Precisión o Exactitud, relacionadas con el rigor metodológico y la adecuación técnica para informar sobre lo que se evalúa; las normas de Responsabilidad para evaluar la calidad de los resultados, desde el equilibrio entre la utilidad, la viabilidad, la adecuación y la precisión.

En la publicación que desarrolla los estándares para la evaluación, cada uno de los estándares se acompaña de una descripción general y una guía para su aplicación, junto con ejemplos para su aplicación, procedentes de diversas situaciones.

El «*Joint Committee*» reconoce y advierte, que los estándares no son igualmente aplicables en todas las evaluaciones. Dependiendo del juicio del profesional de la evaluación, se identificarán aquellos que sean más convenientes de aplicar en cada situación. Los usuarios de los estándares deberían considerar minuciosamente la relevancia de cada estándar para el contexto particular de su aplicación, y decidir cuáles de estos estándares serán más relevantes y cuáles menos. Estas decisiones deberán ser documentadas y argumentadas para su referencia posterior. La evaluación debe facilitar informes basados en evidencias, sobre aspectos positivos y negativos y debe aportar soluciones para mejorar.

Presentamos las 30 normas agrupadas en las cinco categorías o atributos de calidad para las evaluaciones de programas.

4.1. Normas de utilidad

Las normas de utilidad están destinadas a aumentar el grado en que los participantes del programa encuentran los procesos y productos de evaluación valiosos para satisfacer sus necesidades. Una evaluación debe ser útil, para asegurar que sirve a las necesidades de información práctica, centrada en cuestiones importantes. Los estándares de utilidad guiarán la evaluación de tal forma que sea relevante, oportuna y que influya para la mejora.

La utilidad discute el uso, el beneficio, la influencia y el mal uso de la evaluación de un programa. Describe cuándo y cómo se crea el valor de la evaluación, por ejemplo, como cuando las evaluaciones contribuyen al aprendizaje de los grupos interesados, informan las decisiones, mejoran la comprensión, llevan a mejoras o proporcionan información para los juicios de rendición de cuentas.

La utilidad está respaldada por los ocho estándares que se muestran a continuación.

— U1. Credibilidad del evaluador. Las evaluaciones deben ser realizadas por personas cualificadas para establecer y mantener la credibilidad en el contexto de evaluación.

— U2. Atención a las audiencias. Las evaluaciones deben prestar atención a las diferentes personas y grupos que participan o son investigados en el programa y afectados por la evaluación.

— U3. Fines negociados. Los fines de la evaluación deben ser identificados y continuamente negociados basándose en las necesidades de las audiencias.

— U4. Valores explícitos. Las evaluaciones deben aclarar y especificar los valores individuales y culturales que sustentan los propósitos, los procesos y los juicios.

— U5. Información relevante. La información de la evaluación debe servir a las necesidades identificadas y emergentes de las partes interesadas.

— U6. Procesos y productos significativos. Las evaluaciones deben construir las actividades, descripciones y juicios de manera que anime a los participantes a redescubrir, reinterpretar o revisar sus interpretaciones, conocimientos y comportamientos.

— U7. Oportuna y adecuada comunicación y presentación de los informes. Las evaluaciones deben atender a las necesidades de información continua de sus múltiples audiencias.

— U8. Preocupación por las consecuencias y la influencia. Las evaluaciones deben promover el uso responsable y adaptado de los resultados de la evaluación tomando medidas de protección ante las consecuencias negativas del posible uso indebido y/o alguna perdida no intencionada.

4.2. Normas de viabilidad o factibilidad

Las normas de viabilidad o factibilidad están dirigidas a incrementar la efectividad y eficiencia de la evaluación.

La factibilidad discute los efectos de contextos, culturas, costes, políticas, poder, recursos disponibles y otros factores en las evaluaciones. Detalla los factores de factibilidad a considerar antes de implementar una evaluación y cómo aumentar los factores a considerar antes de implementar una evaluación y cómo aumentar o mantener la factibilidad en diferentes contextos.

Una evaluación debe estar diseñada y conducida de tal modo que sea realista, prudente, diplomática, ajustada y frugal en relación con el coste-beneficio. Una evaluación debe ser viable, factible, y asegurar que es realista y eficiente. Debe utilizar procedimientos eficientes que no ocasionen demasiados problemas. Los estándares de viabilidad reconocen que las evaluaciones se realizan en contextos naturales, sus diseños deben ser operativos y sus gastos, en material, personal o tiempo, no deben exceder al requerido por las cuestiones de la evaluación.

La factibilidad es apoyada por cuatro estándares.

— F1. Gestión de proyectos. Las evaluaciones deben utilizar estrategias efectivas de gestión de proyectos.

— F2. Procedimientos prácticos. Los procedimientos de la evaluación deben ser prácticos y sensibles a la forma en que funciona el programa.

— F3. Viabilidad contextual. Las evaluaciones deben reconocer, controlar y equilibrar los intereses culturales y políticos y las necesidades de los individuos y de los grupos.

— F4. Uso de recursos. Las evaluaciones deben utilizar los recursos de manera eficaz y eficiente.

4.3. Normas de probidad o adecuación legal y ética

Probidad tiene que ver con las preocupaciones morales, éticas y legales relacionadas con la calidad de la evaluación. Las normas de probidad apoyan lo que es propio, adecuado, legal, correcto, honrado, íntegro, recto y justo en las evaluaciones. Una evaluación debe asegurar que es conducida de modo legal y éticamente, basada en compromisos explícitos, que aseguren la necesaria cooperación, la protección de los grupos implicados y la honestidad de los resultados.

Las normas de probidad, o legitimidad, consideran los derechos de las partes interesadas y otras personas y detallan las responsabilidades de todas las partes interesadas, especialmente de los profesionales de la evaluación, en una evaluación. Establecer umbrales para una probidad adecuada puede ser difícil y requiere equilibrar las necesidades y situaciones de los diferentes actores. Las consideraciones de justicia social pueden desempeñar un papel importante, pero no todas las partes interesadas tienen los mismos valores o concepciones de la justicia social.

La calidad de las evaluaciones en relación con la adecuación, probidad o legitimidad, está respaldada por siete estándares.

— P1. Orientación de confianza e integradora. Las evaluaciones deben responder a los interesados y sus comunidades.

— P2. Acuerdos formales. Los acuerdos de la evaluación deben ser negociados para hacer explícitas las obligaciones y tener en cuenta las necesidades, expectativas y contextos culturales de los clientes y otros grupos interesados.

— P3. Derechos humanos y respeto. Las evaluaciones deben ser diseñadas y realizadas para proteger los derechos humanos y legales y mantener la dignidad de los participantes y de las audiencias (los interesados).

— P4. Claridad y equidad. Las evaluaciones deben ser comprensibles y justas para atender las necesidades y fines de los interesados.

— P5. Transparencia y divulgación. Las evaluaciones deben facilitar una descripción completa de los objetivos, las limitaciones y las conclusiones a todos los interesados a menos que el hacerlo violara las obligaciones legales.

— P6. Conflictos de intereses. Las evaluaciones deben ser abiertas y honestamente identificar y abordar los conflictos de intereses reales o percibidos que puedan comprometer la evaluación.

— P7. Responsabilidad fiscal. Las evaluaciones deben calcular todos los recursos asignados y cumplir con los procedimientos y procesos fiscales.

4.4. Normas de precisión o exactitud

Las normas de exactitud están destinadas a aumentar la fiabilidad y veracidad de las representaciones de la evaluación, de sus proposiciones y hallazgos, especialmente aquellos que apoyan las interpretaciones y los juicios sobre la calidad. Las normas de exactitud o precisión se preocupan de cómo aumentar la precisión de los hallazgos y las conclusiones. La precisión discute la fiabilidad, validez y reducción del error y sesgo. Los ocho estándares de precisión tratan la calidad en la recogida de datos, su análisis, lógica y coherencia, sus conclusiones y la comunicación de sus procedimientos y hallazgos.

Una evaluación debe ser exacta, para asegurar que revela y comunica la información encontrada, describe con claridad el objeto evaluado en

su evolución y en su contexto, revela las virtudes y defectos del plan de evaluación, de los procedimientos y de las conclusiones y proporciona conclusiones válidas y fidedignas. Las conclusiones y los juicios deben ser coherentes con los datos.

Se espera que las normas de exactitud aseguren que la evaluación revelará y aportará la información técnicamente adecuada sobre los rasgos que determinan el valor o mérito del programa evaluado. Las normas de exactitud están respaldadas por ocho estándares.

— A1. Conclusiones y decisiones justificadas. Las conclusiones y decisiones de la evaluación deben estar justificadas explícitamente en las culturas y contextos en los que tienen consecuencias.

— A2. Información válida. La información de la evaluación debe responder a los fines previstos y ser la base de interpretaciones válidas.

— A3. información fiable. Los procedimientos de evaluación deben proporcionar información lo suficientemente confiable y consistente para los usos previstos.

— A4. Descripciones explícitas del programa y del contexto. Las evaluaciones deben documentar los programas y sus contextos con el detalle y amplitud adecuados a los propósitos de evaluación.

— A5. Gestión de la Información. Las evaluaciones deben emplear métodos sistemáticos de recogida, revisión, verificación y almacenamiento de la información.

— A6. Diseños y análisis adecuados. Las evaluaciones deben emplear diseños y análisis técnicamente adecuados y que sean apropiados para los fines de la evaluación.

— A7. Razonamiento explícito de la evaluación. El razonamiento de la evaluación que desde la información recogida y su análisis conduce a los resultados, las interpretaciones, las conclusiones y los juicios debe ser claramente y completamente documentado.

— A8. Comunicación y presentación de los informes. Las comunicaciones de la evaluación deben tener la perspectiva adecuada y protegerse contra los conceptos erróneos, los prejuicios, las distorsiones y los errores.

4.5. Normas de responsabilidad de la evaluación

La responsabilidad de la evaluación abarca el atributo de la calidad de la evaluación como resultado del equilibrio entre los atributos de utilidad, viabilidad, legitimidad y exactitud. La metaevaluación interna y externa, que integran las normas de rendición de cuentas, y que deben ser aplicadas de un modo completo, proporcionan la metodología para incrementar y documentar la calidad de la evaluación.

La responsabilidad de la evaluación introduce tres nuevas normas, pero también está respaldada por las 27 normas específicas de utilidad, factibilidad, propiedad y exactitud.

Las normas de rendición de cuentas o de responsabilidad de la evaluación fomentan la completa, correcta y adecuada documentación de las evaluaciones, junto con una perspectiva metaevaluativa enfocada a la mejora y la rendición de cuentas, de los procesos y de los resultados de la evaluación.

— E1. Documentación de la evaluación. Las evaluaciones deberán documentar plenamente sus propósitos negociados y los diseños, procedimientos, datos y resultados aplicados.

— E2. Metaevaluación interna. Los evaluadores deben utilizar estas y otras normas aplicables para rendir cuentas del diseño, los procedimientos empleados, la información recopilada y los resultados obtenidos de la evaluación.

— E3. Metaevaluación externa. Los patrocinadores, los clientes, los evaluadores y otras partes interesadas en la evaluación del programa deberían alentar la realización de metaevaluaciones externas usando estas y otras normas aplicables.

> El objetivo de la metaevaluación es asegurar la calidad de los servicios evaluativos, evitar enfrentarse a prácticas ilegales o a servicios que no son de interés público, señalar el camino para el perfeccionamiento de la profesión y promover una mayor comprensión de la empresa evaluativa. (Stufflebeam 1987: 52).

5. UTILIZACIÓN DE LAS NORMAS DEL JOINT COMMITTEE PARA EL DESARROLLO DE LOS PROGRAMAS, DE SU EVALUACIÓN Y DE SU RESPONSABILIDAD PÚBLICA

El orden de presentación de los atributos o características de calidad y sus estándares para la evaluación de programas no indica su mayor o menor importancia. Todos deben ser utilizados de modo efectivo, adaptativo, sensible y con significado de su utilización. Históricamente se creía que los estándares de precisión eran los más importantes para realizar una evaluación de calidad. Sin embargo, el énfasis posterior sobre los atributos de utilidad de las evaluaciones han reconocido que una evaluación muy precisa pero inútil tiene poco o nada que ver con su valor. Por esto las normas de utilidad vienen en primer lugar desde la primera edición de las normas, si bien no puede ser la única o principal característica de una evaluación de calidad.

Los otros atributos de calidad de las evaluaciones de programas son igualmente críticos. Así, ignorar la viabilidad de la evaluación puede resultar en un despilfarro de recursos mediante una realización fallida que pueda alterar las necesidades de evaluación del programa. (Yarbrough, Shulha, Hopson & Caruthers, 2011: xxiii).

El atributo de probidad y legalidad, enfatiza la primacía de los derechos humanos y la responsabilidad del trabajo en el propio programa y en su evaluación. La desatención de este tipo de normas pude tener serias consecuencias éticas, morales, profesionales y legales.

La atención a las normas de precisión, o exactitud, proporciona la fundamentación en los hechos requerida para que las evaluaciones sean útiles. La responsabilidad en la evaluación es un requerimiento global, dominante, debido a su énfasis en la naturaleza autoreflexiva de la calidad de la evaluación y la necesidad de investigar y documentar el equilibrio entre los otros atributos de calidad en cada situación de la evaluación como es realizada toda evaluación en tiempo y espacio real.

Utilizar los estándares de modo efectivo requiere comprenderlos en profundidad. De otro modo se perdería la riqueza que hace los estándares útiles. Considerados en su totalidad, los estándares para la evalua-

ción de programas se centran en las dimensiones de calidad para la evaluación de programas. Las dimensiones de calidad tienen que ver con la utilidad de la evaluación, su viabilidad, su adecuación, su precisión y su responsabilidad. Tomadas en conjunto, la aplicación de las normas, cuentan la historia de cómo se construye una evaluación de calidad en situaciones específicas.

Los evaluadores prácticos tienen especial responsabilidad en la aplicación de las normas, debido a que ellos son responsables sobre muchas de las característica para que una evaluación sea de calidad. Por ejemplo, los profesionales de la evaluación pueden incrementar su calidad si al comenzar una evaluación:

— Se comunican con sus clientes, y juntos valoran si el programa es evaluable.

— Informan a sus clientes acerca de si se puede o no hacer una evaluación.

— Identifican e interactúan con los implicados.

— Planifican la evaluación, con frecuencia, con los implicados y posibles beneficiarios de la evaluación.

— Organizan la realización de la evaluación, incluyendo la recogida de información y su análisis, su comunicación e informes y seguimiento de los hallazgos.

— Investigan, documentan y mejoran la calidad de la evaluación y la información de la misma para el rendimiento de cuentas a través de la realización de la metaevaluación.

— Realizan y difunden informes sobre las evaluaciones y el uso apropiado de las mismas. (Yarbrough, Shulha, Hopson & Caruthers, 2011: xxxiv).

La orientación hacia el uso responsable de los estándares de calidad se relaciona con los programas y con la evaluación de los mismos. De ese modo, los estándares deben garantizar de un modo responsable, tanto en relación con los programas como en relación con su evaluación:

— El uso adecuado de los recursos.

— La realización efectiva de las metas pretendidas de acuerdo con las necesidades de los beneficiarios del programa.

— Todas las consideraciones de adecuación ética y legal.

— Investigar sobre la mejora de las necesidades.

— Responsabilizarse de aplicar las mejoras paso a paso.

— El alineamiento de las especificaciones para atender a las necesidades, los procesos necesarios para ello y los resultados para atender las necesidades. (Ibíd.: xxxvii).

Un ejemplo de la aplicación de las Normas puede encontrarse en el artículo de Chatterji (2005) que se cita en las referencias.

6. LISTA DE CONTROL DE LAS NORMAS PARA LA COMPROBACIÓN DE SU APLICACIÓN EN LA EVALUACIÓN DE PROGRAMAS

La lista de control conteniendo los estándares de calidad en los cinco atributos o características que los conforman, facilita su aplicación. La publicación aporta una 'Lista de comprobación para la aplicación de los estándares' pero advierte de que su aplicación de una manera simplificada no sería la mejor manera de promover la mejora de la calidad de los programas y de la educación. Incluimos la tabla con la lista de los estándares para la aplicación y comprobación de su aplicación.

El estándar puede recibir la valoración de 1 a 4, de acuerdo con el siguiente significado:

1. El estándar no era aplicable.
2. El estándar no se utilizó.
3. El estándar se utilizó parcialmente.
4. El estándar se utilizó.

Tabla 1. La lista para la comprobación de la aplicación de los estándares

LISTA PARA LA COMPROBACIÓN DE LA APLICACIÓN DE LOS ESTÁNDARES					
		1	**2**	**3**	**4**
Utilidad	U1 Credibilidad del evaluador				
	U2 Atención a las audiencias				
	U3 Fines negociados				
	U4 Valores explícitos				
	U5 Información relevante				
	U6 Procesos y productos significativos				
	U7 Oportuna y adecuada comunicación y presentación de los informes				
	U8 Preocupación por las consecuencias y la influencia				
Factibilidad	F1 Gestión de proyectos				
	F2 Procedimientos prácticos				
	F3 Viabilidad contextual				
	F4 Uso de recursos				
Legitimidad	P1 Orientación de confianza e integradora				
	P2 Acuerdos formales				
	P3 Derechos humanos y respeto				
	P4 Claridad y equidad				
	P5 Transparencia y divulgación				
	P6 Conflictos de intereses				
	P7 de Responsabilidad fiscal				

LISTA PARA LA COMPROBACIÓN DE LA APLICACIÓN DE LOS ESTÁNDARES					
		1	2	3	4
Exactitud	A1 Conclusiones y decisiones justificadas				
	A2 Información válida				
	A3 Información fiable				
	A4 Descripciones explícitas del programa y del contexto				
	A5 Gestión de la información				
	A6 Diseños y análisis adecuados				
	A7 Razonamiento explícito de la evaluación				
	A8 Comunicación y presentación de los informes				
Responsabilidad	E1 Documentación de la evaluación				
	E2 Metaevaluación interna				
	E3 Metaevaluación externa				

— Si contesta el valor 1, explique, pruebe, por qué el estándar no era aplicable.

— Si contesta el valor 2, explique por qué el estándar no se utilizó.

— Si contesta el valor 3, explique por qué el estándar se utilizó parcialmente.

La utilización de la lista de control para la comprobación de la aplicación de los estándares para la evaluación de programas, mediante la tabla anterior, requiere que se dé una explicación del por qué una determinada norma no ha sido aplicada.

Bajo cada uno de los cinco atributos de la calidad de la evaluación y las 30 normas que los integran, se incluye una diversidad de líneas detalladas para orientar su aplicación, que pueden consultarse en el texto completo de los Estándares de 2011.

Las normas del *Joint Committee* aportan criterios a tener en cuenta para la realización de evaluaciones y para la valoración de las mismas. Una evaluación que no cumpla dichos requisitos puede crear más perjuicios que beneficios por lo que no sería aconsejable hacerla. Estas normas sirven de guía para la realización de evaluaciones con el fin de que sean útiles, factibles, éticas y exactas, requisitos fundamentales para que la evaluación cumpla con su objetivo primordial: la toma de decisiones informada que permita introducir mejoras ajustadas a las necesidades reales que motivaron el diseño del programa y las nuevas necesidades detectadas.

La aplicación de estos cinco grupos de normas asegura:

— *Una evaluación integral* que afecta a la organización y a las personas implicadas en el programa, desde sus metas y objetivos, hasta la planificación, estrategias, secuenciación y recursos para conseguirlos, los resultados y sus conclusiones y valoraciones.

— *Una evaluación sistemática*, estructurada siguiendo un plan consensuado, asegurando la objetividad y honradez de los resultados.

— *Una evaluación innovadora,* que permite ir comprobando el progreso, así como los elementos que inciden en la mejora continua del programa.

— *Una evaluación para la mejora*, que al ser exacta por fundamentarse en evidencias contrastadas, permite establecer planes de mejora a corto y medio plazo y la valoración progresiva de los avances.

— *Una evaluación comprometida con la responsabilidad,* que al comunicar con exactitud sus hallazgos, contribuye a la mejora de los programas, de su evaluación y a la mejora de la propia disciplina de Evaluación de Programas.

EJERCICIOS DE AUTOEVALUACIÓN

1. Haga una valoración de cada uno de los grupos de normas para juzgar la calidad de las evaluaciones y aporte argumentos.

2. Evalúe un programa con el que esté familiarizado utilizando los cinco grupos de atributos o características de calidad de las evaluaciones, y sus correspondientes normas del Joint Committee, y comente el grado de aplicación de los estándares y explique si alguno no pudo aplicarse a qué se debió.
3. Valore las posibilidades de aplicación de los estándares para la mejora de los programas, su evaluación y la responsabilidad o rendimiento de cuentas.

REFERENCIAS BIBLIOGRÁFICAS

Chatterji, M. (2005). «Applying the Joint Committee's 1994 Standards in International Contexts: A case Study of Educational Evolutions in Bangladesh». Teachers College Record, 107, 10, 2372-2400.

Joint Committee on Standards for Educational Evaluation (1981). *Standards for evaluation of educational programs, projects and materials.* New York: McGraw-Hill.

— (1994). «The Program Evaluation Standards». Thousand Oaks, CA: Sage P. Inc.

Program Evaluation Standards Statements. Accesible en http://www.jcsee.org/program-evaluation-standards-statements

Sanders, J. R. (Cord.) (1998). «Estándares para la evaluación de programas». *Comité Conjunto de estándares para la evaluación educativa,* Mensajero. Bilbao, 1998.

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., and Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.

Tema 7

Técnicas e instrumentos para la recogida de información en la evaluación de programas

1. Objetivos y resultados de aprendizaje.
2. Introducción.
3. La observación.
   - 3.1. La observación sistemática y la observación participante.
   - 3.2. Los sistemas de recogida de información en la observación.
   - 3.3. Cuestiones sobre el muestreo y las fuentes de error en la observación.
   - 3.4. Instrumentos para el registro de la observación.
   - 3.5. Características de la observación participante.
     - 3.5.1. La función del observador en la observación participante.
     - 3.5.2. El diseño para la realización de la observación participante.
     - 3.5.3. El proceso de recogida de datos en la observación participante
     - 3.5.4. El análisis de la información en la observación del participante.
     - 3.5.5. La elaboración del informe de la observación participante.
4. La evaluación de necesidades.
5. El cuestionario.
6. La entrevista.
7. El análisis de contenido.
   - 7.1. Las fases del análisis de contenido.
   - 7.2. Las unidades de análisis y su contexto.
   - 7.3. La inferencia como objetivo intelectual del análisis de contenido y la validez como criterio de éxito.

7.4. La inferencia como objetivo intelectual del análisis de contenido y la validez como criterio de éxito.

8. La técnica Delphi.

9. Los grupos de discusión.

Ejercicios de autoevaluación.

Referencias bibliográficas.

1. OBJETIVOS Y RESULTADOS DE APRENDIZAJE

Se espera que el alumno, tras el estudio de este tema, sea capaz de identificar la técnica de recogida de información y datos más adecuada a los objetivos de la evaluación que va a realizar, así como los procedimientos para su análisis y la elaboración de conclusiones de acuerdo con los resultados obtenidos en el proceso de realización de la evaluación de un programa educativo. De un modo más concreto pretende servir al logro de los siguientes objetivos de aprendizaje:

— Revisar algunos de los procedimientos para la recogida de información más utilizados en Educación.

— Valorar la adecuación de cada uno de los procedimientos para la recogida de información de acuerdo con los objetivos de la evaluación de un programa, los momentos, las situaciones y ámbitos en los que se aplica.

— Elaborar instrumentos de recogida de información adaptados a las diferentes necesidades de evaluación de un programa.

2. INTRODUCCIÓN

Este tema es fundamental para la realización de evaluaciones de programas reales. Complementa lo presentado en los temas referidos a los modelos de evaluación de programas, incluido el referido a los procedimientos para la evaluación de programas siguiendo los modelos lógicos. En estos temas se menciona la evaluación de necesidades para conocer las condiciones de las personas en sus contextos, sobre las cuales se va a

intervenir mediante la aplicación de un proyecto, programa o intervención, y muestra la conveniencia de conocer el punto de partida, las necesidades y problemas. Para ello es necesario realizar observaciones, encuestas, entrevistas. También es necesario realizar análisis de informes de resultados de investigaciones previas para comparar lo investigado con la situación sobre la que debemos intervenir que nos ayudara a elaborar las nuevas propuestas de mejora de la situación a atender. Y una vez diseñada nuestra intervención, considerando el propio programa, sus objetivos y estrategias de aplicación y sus metas y resultados pretendidos, junto con los recursos necesarios, deberemos analizar si se ajusta a las necesidades, y si se orienta a conseguir resultados, metas, de calidad. Y deberemos recoger información sobre los procesos de aplicación del programa mediante observación y cuestionarios, pruebas de rendimientos, análisis de los resultados que se van consiguiendo mediante grupos de discusión con representantes clave. Y analizar los resultados conseguidos para valorar la eficacia del programa y de su aplicación, y la satisfacción en sus usuarios y beneficiarios, y asimismo podría verse conveniente realizar innovaciones a la luz de nuevas necesidades detectadas para lo que podría ser conveniente realizar un Delphi con expertos.

Para poder realizar este tipo de acciones, analizar situaciones, planificar intervenciones, diseñar programas y recoger información, analizarla y valorarla y elaborar informes de todo ello, es necesario profundizar en las diferentes técnicas de recogida y análisis de información y datos. Y ese es el objetivo de este tema, entendiendo que estos contenidos, junto con el análisis de datos, han sido ya objeto de estudio específico en las asignaturas de Metodología de Investigación ya estudiadas por el alumno en cursos previos a éste en el Grado.

3. LA OBSERVACIÓN

La observación es una de las técnicas de investigación más importante en las Ciencias Sociales y, por tanto, para su aplicación en la evaluación, ya que nada puede reemplazar el contacto directo del observador investigador con las personas en las situaciones de estudio. Cualquier procedimiento para el estudio y comprensión de la realidad necesita de una primera fase de observación que permite tener una idea real de lo que se va a estudiar.

La observación es un proceso intelectual que requiere atención, concentración selectiva, guiada por objetivos, que ayuda a organizar lo observado en una síntesis integrada en una estructura con significado. La observación supone contemplar de un modo sistemático, cómo se desarrolla aquello que queremos estudiar, sin manipularlo, tal cual discurre, siendo la función del observador recoger información de aquello que observa. La observación supone la recogida de información guiada por objetivos: observar para valorar unas necesidades, observar para establecer un diagnóstico, observar para confirmar una hipótesis de investigación, observar para evaluar unos resultados. (En Postic y De Ketele, 1988:17).

La observación puede utilizar al observador como instrumento de recogida de datos, o puede servirse de una pauta, o guion de observación, lo que sería observación en sentido estricto, o de un cuestionario de encuesta, en sentido más amplio. La observación supone las operaciones de selección, estructuración e interpretación de la información observada, de modo que quede patente una red de significados.

La observación se considera un proceso contextualizado, en el sentido localista, o del lugar en el que se producen los acontecimientos a observar y los contextos que lo rodean, en un aula, centro, comunidad, y en el sentido temporal e histórico de las situaciones, incluyendo las tradiciones, los estilos de comunicación, las expectativas de los protagonistas, y de los acontecimientos, su devenir histórico mediante el estudio de memorias, informes, actas, investigaciones previas, y todo ello ayuda al investigador a comprender los fenómenos que se producen en el contexto que se estudia. (Evertson y Green, 1997:307).

3.1. La observación sistemática y la observación participante

Son dos los tipos de observación denominadas sistemática y participante, en función de la mayor o menor predeterminación de las categorías a observar, el modo de aproximación al contexto, los instrumentos que utiliza y los procedimientos y funciones del observador.

En la observación sistemática las categorías que guían la observación están previamente establecidas, basándose en los objetivos a observar y las investigaciones previas. En este tipo de observación la característica

principal del observador es la ausencia de interacción con lo que observa; el observador contempla lo que está aconteciendo de modo distanciado, y registra los hechos en una hoja marcando la presencia o ausencia de las categoría preestablecidas.

En el extremo opuesto, en la observación participante, el observador se introduce en la situación a observar sin partir de unas premisas preestablecidas, aunque hay que decir que el observador participante no va con su mente en blanco, vacía, sino que también tiene unos objetivos fundamentados en la teoría y conocimientos previos, aunque trata de no condicionar lo que observa, esperando que emerja durante el tiempo de su permanencia en el campo, tratando de ser uno más en la situación en la que está inmerso y que estudia, lo que le permite obtener información que de otro modo no podría conseguir, incluida la información procedente de su propia experiencia subjetiva; toma notas en su cuaderno de campo, que completará inmediatamente a su salida del mismo, para mantener más cercano el recuerdo.

Ambos enfoques de observación son aplicables en la evaluación de programas para analizar los contextos y realizar análisis de necesidades antes de la aplicación de un programa Para analizar los procesos durante la aplicación de un programa. Para recoger información en las reuniones de grupos, de expertos, sobre los progresos de los programas y en reuniones de padres de alumnos, entre otras muchas situaciones en las que la utilización de las metodología de observación son aconsejables.

3.2. Los sistemas de recogida de información en la observación

Podemos considerar cuatro sistemas de recogida de información mediante la observación, en función de los principios en los que se fundamentan: categoriales, descriptivos, narrativos y tecnológicos.

— *El sistema categorial* pretende recoger información sobre un limitado número de categorías o conductas. Estas unidades de observación se elaboran deductivamente a partir de la teoría, a partir de investigaciones previas. Son unidades discretas, reflejando cada categoría una sola conducta que puede ser considerada una variable independiente. El observador recoge información sólo de los ítems previamente seleccionados y utiliza dos tipos de

límites en el período de observación, el del tiempo y el de los acontecimientos o conductas a observar. Los instrumentos para recoger la información pueden ser inventarios o listas de comprobación de la aparición de las conductas seleccionadas, o escalas, por ejemplo, tipo Likert, o el diferencial semántico de Osgood, indicando su intensidad y frecuencia, o número de veces que se presenta la conducta en el período de tiempo seleccionado. Los resultados de las observaciones, registradas en términos numéricos, permiten la realización de análisis estadísticos. (Ibíd.:174).

— *El sistema descriptivo* no contiene un número limitado de categorías, aunque puede trabajar con un rango de categorías fijas para describir los comportamientos, pudiendo generarse éstas a partir de los modelos observados, o haber sido identificadas a partir de datos previos. Las unidades de observación que utilizan los sistemas descriptivos tienen una base conceptual. El foco del observador es amplio para captar la mayor amplitud de lo que acontece en la situación de estudio; la información es recogida dentro de unos márgenes de acontecimientos y contextos sobre las conductas a observar, transcribiéndolas con lenguaje verbal, utilizando medios manuales o técnicos para realizar grabaciones. La meta del investigador es obtener una descripción detallada del fenómeno para explicar el desarrollo de los procesos y tratar de identificar principios generales.

— *El sistema narrativo* es un sistema abierto por no trabajar con categorías prefijadas. Se basa en el estudio del lenguaje oral o escrito para recoger las descripciones sobre las conductas que estudia a través de los diarios o anecdotarios del profesor, o de las notas de campo del observador participante. Las unidades de observación en los sistemas narrativos son similares a las de los sistemas descriptivos pero sobre información textual. El muestreo de los acontecimientos ocurre dentro de sus límites naturales. Los datos están influidos por la percepción, el marco de trabajo, la experiencia y la riqueza lingüística del observador que narra lo que observa. La meta del investigador es explorar, describir y explicar el desarrollo de los procesos e identificar principios generales y específicos, como en el sistema descriptivo, buscando, además, identificar y comprender los factores que influyen en la ocurrencia de dichos acontecimientos. Los hallazgos encontrados son útiles para la comprensión de casos similares.

— *El sistema tecnológico* utiliza medios como las videocámaras para las grabaciones de los acontecimientos que proporcionan la materia pri-

ma, permitiendo seleccionar objetivos para su estudio «a posteriori» mediante cualquiera de los sistemas mencionados. El muestreo de los acontecimientos ocurre dentro de sus límites naturales, de igual modo que en los sistemas anteriores. Las grabaciones facilitan al investigador y a los participantes el recuerdo para obtener las perspectivas de éstos sobre los acontecimientos y validar los modelos observados en una forma de triangulación entre observadores, y de confirmación, por la posibilidad de repetir la visualización o audición cuantas veces sea necesaria. La meta del investigador que utiliza medios tecnológicos es obtener un registro de un acontecimiento o fenómeno para poder estudiarlo en profundidad en un momento posterior. Los medios tecnológicos pueden utilizarse para complementar los otros sistemas mencionados.

3.3. Cuestiones sobre el muestreo y las fuentes de error en la observación

Seleccionar el momento, el lugar, el tiempo y los centros de observación asociados a la cuestión de estudio son aspectos de fundamental importancia relacionados con las cuestiones del muestreo que deben estar guiados por el marco teórico del estudio que se realiza (Evertson y Green, 1986:180). Son cuatro las áreas de decisiones sobre el muestreo:

— El número de sujetos.

— La duración total del tiempo de observación.

— La distribución del tiempo de observación.

— La probabilidad de que la observación sea representativa del fenómeno objeto de estudio (qué se observa, cuándo y dónde).

En el sistema categorial las decisiones sobre el muestreo se adoptan antes de comenzar la observación, de un modo apriorístico. En los sistemas descriptivo, narrativo y tecnológico se adoptan a lo largo del proceso de investigación.

El muestreo del tiempo supone considerar los períodos de observación y los intervalos de grabación si fuera necesario. Para el muestreo de los acontecimientos, el investigador determinará el asunto específico que

será el objeto de la observación y determinará las situaciones posibles en las que el fenómeno puede ocurrir y los centros de observación donde pueden manifestarse de un modo representativo, para garantizar la fiabilidad y validez de la observación.

Las fuentes de error tienen relación con el observador, el sistema de representación elegido, el marco conceptual del fenómeno estudiado y los procedimientos de recogida de datos, y afectan a la fiabilidad, la validez y al rigor de las observaciones. Los errores van unidos al sistema perceptivo, lo que lleva a los investigadores a adoptar variedad de medidas para reducir el error, asegurar el rigor y obtener una representación adecuada de la realidad mediante la observación.

La validez se refiere al grado en que la información obtenida por un instrumento realmente describe lo que el investigador se había propuesto describir y la fiabilidad se refiere a la exactitud y consistencia con que lo hace; para considerar la fiabilidad en las investigaciones observacionales hay que hacer dos distinciones conceptuales entre los índices de fiabilidad procedentes de los coeficientes de fiabilidad utilizados para controlarla mediante, por ejemplo, registros paralelos, o el procedimiento de recogida de información sucesiva, a cuyos resultados numéricos se les aplica pruebas estadísticas para hallar coeficientes de fiabilidad, como los utilizados en las pruebas de fiabilidad de las escalas de Likert. También se pueden hallar indicadores de fiabilidad mediante la utilización de coeficientes de acuerdo entre varios observadores sobre los acontecimientos que observan y describen. Cuando el registro de las observaciones se realiza mediante medios técnicos, el visionado repetido de las grabaciones por varios observadores, garantizará la fiabilidad de las interpretaciones y de los resultados de la observación. (Ibíd.:184).

La fiabilidad, como la validez de los datos, se verán favorecidas si se cuida en los instrumentos de recogida de datos la claridad en la formulación de las definiciones conductuales a observar, el número de conductas o categorías a observar, que deberá ser reducido, y el muestreo de tiempo que deberá ser suficiente para garantizar el registro emitido.

Todas estas cuestiones contribuyen a la validez de las investigaciones en las ciencias sociales, incluida la investigación evaluativa.

3.4. Instrumentos para el registro de la observación

Los instrumentos para el registro de observaciones se clasifican, de acuerdo con su mayor o menor estructuración, como siguen:

— *Registros de conducta, mediante tablas de doble entrada*, en las cuales en las columnas se especifican las conductas y en las filas se indica la ocurrencia o no de esa conducta y su intensidad durante el período temporal de observación fijado. Es una forma predeterminada de registro, dado que las conductas se especifican antes de que sucedan. Pueden utilizarse para registrar la aparición de determinadas conductas en un sujeto (Figura 1) o para estudiar una conducta en el grupo.

— *Escalas de valoración, o inventarios de conducta*, en los cuales las conductas seleccionadas son valoradas mediante una escala de grados, por ejemplo, de 1 a 5 de menor a mayor aparición de la conducta especificada.

— *Registros narrativos abiertos en contextos naturales*, en los que el observador toma notas sobre lo que acontece en la situación que observa, que posteriormente completará con sus apreciaciones.

Ejemplo de registro individual:

REGISTRO INDIVIDUAL			
Nombre:			
Conductas	**Grado de aparición**		
	Ausente	**Rara vez**	**Frecuente**
Atiende las intervenciones de los demás			
Colabora en las actividades			
Sugiere nuevas actividades			
Disfruta en las actividades grupales			
Se mantiene aislado…			

Figura 1. Ejemplo de registro individual

Ejemplo de registro narrativo abierto. El registro anecdótico sobre un acontecimiento crítico es importante anotarlo en el momento más cercano a su ocurrencia para evitar olvidar información relevante (Figura 2). Es igualmente importante diferenciar la descripción de lo acontecido de la interpretación del mismo. Un registro anecdótico es un relato verbal que muestra las siguientes características:

1. Cada anécdota debe limitarse a un solo incidente.
2. Debe contener una descripción lo más literal posible de lo observado. Por ejemplo, El estudiante dijo: «He disfrutado como nunca utilizando el ordenador para aprender, en lugar de El estudiante expresó su satisfacción con el sistema de enseñanza».
3. Debe incluir una descripción de la situación en la que el incidente ocurre para que el significado de la conducta pueda entenderse.
4. Debe escribirse lo más pronto posible después de presenciar u oír hablar del incidente para que puedan incluirse todos los detalles importantes.
5. Debe incluir una sección, separada, que describa la interpretación sobre el incidente anecdótico. La valoración personal, como experto de lo que observa, es importante.

HOJA DE REGISTRO ANECDÓTICO
Nombre de observador: ..... Fecha:..... Lugar.....
Descripción del incidente:
Situación en la que se da el incidente:
Interpretación:

Figura 2. Ejemplo de registro anecdótico.

Ejemplo de registro para la revisión del proceso de aplicación de una innovación. Algunas innovaciones pueden fracasar por no haber considerado algunos factores de la aplicación, como los recursos necesarios o la motivación de quienes las van a aplicar. Para comprobar cómo se están aplicando las innovaciones es esencial recoger información relevante acerca del funcionamiento real del programa, comparándolo con lo pla-

neado. La herramienta para el *Registro de aplicación del programa* (Figura 3) se ha diseñado para hacer esa comparación de un modo sistemático. Mostramos unas breves indicaciones para su realización.

Además de los encabezamientos para identificar el programa y cualquier otra información específica, un registro sobre la implementación de un programa debería incluir lo siguiente: una columna para describir lo que se planeó que ocurriera durante la aplicación, otra para describir lo que realmente ocurrió, otra para comentar las diferencias entre lo planeado y lo ocurrido, y un espacio adicional para comentarios importantes sobre el contexto.

HOJA DE REGISTRO SOBRE LA IMPLEMENTACIÓN DE UN PROGRAMA

Nombre del programa: .....

Profesor: ..... Número de asistentes: .....

Fecha:..... Lugar.....

Horario	Actividades planeadas	Actividades reales	Comentarios
09:00-09:15	Introducción de los participantes y revisión de la agenda (Profesor)	Ningún cambio	
09:15-09:30	Presentación global del «Nuevo Curso» (Profesor)	Ningún cambio	
09:30-11:30	Los estudiantes comienzan el Módulo 1 del curso	El sistema falló porque alguien borró los archivos	
11:30-12:00	Descanso	Dos estudiantes decidieron seguir trabajando	Necesito verificar todos los sistemas antes de que comience el curso
12:00-14:00	Los estudiantes continúan trabajando a través de los módulos	Ningún cambio	Tuve que ayudar a los estudiantes en el aprendizaje del módulo con frecuencia

1. ¿Qué actividades de formación le gustaría modificar la próxima vez que dirija este curso?
Verificaré antes de comienzo de la actividad cada uno de los sistemas multimedia para asegurar que los módulos funcionan tal como se diseñaron.

2. ¿Cómo pueden utilizarse los materiales de formación en este curso para reforzarlo?
Los estudiantes señalaron varios errores en el módulo de datos. Quedan recogidos en la lista adjunta.

Figura 3. Ejemplo de registro sobre la implementación de un programa.

3.5. Características de la observación participante

La observación participante está en el extremo del continuo menos estructurado de la observación. La observación participante es una estrategia metodológica utilizada en el trabajo de campo, desarrollada en la etnografía y en la antropología. Combina el uso de varias técnicas como son el análisis de documentos, la entrevista a grupos y personas considerados como informantes claves, la participación directa del observador en el campo y la introspección.

La utilización de los métodos de la etnografía se ha mostrado eficaz en los ámbitos educativos y sociales por su carácter descriptivo, exploratorio e interpretativo, de acuerdo con la idea de que en la interacción con los demás existe siempre el objetivo de obtener información sobre las personas en las situaciones en las que viven y se desarrollan.

3.5.1. La función del observador en la observación participante

En la observación participante el investigador se integra tanto como pueda, en las situaciones que estudia. El propósito de tal participación es desarrollar un punto de vista interno, del propio contexto, acerca de lo que está ocurriendo. Sin embargo, el profesional de la observación tiene que combinar el lado del observador y el de participante, de modo que sea capaz de comprender la situación desde un punto de vista interno y a la vez aportar la perspectiva del profesional externo. Es conveniente que el observador mantenga una cierta distancia profesional con el fin de favorecer la objetividad de la observación. Entre el observador y el grupo observado, se establece una comunicación, normalmente de carácter no verbal, en la que el observador está alerta a las claves de la realidad que observa para, a través de ellas, interpretar lo que ocurre de un modo profundo.

Desde la observación participante se defiende el análisis desde la perspectiva de los propios actores del grupo estudiado, tratando de captar sus puntos de vista. Sin embargo, el observador, debido a su esfuerzo consciente de objetividad para observar los acontecimientos y los cam-

bios que allí se produzcan, alcanza, en cierta medida, una postura marginal, para captar dichos cambios.

3.5.2. El diseño para la realización de la observación participante

El proceso de investigación mediante la observación participante conlleva las siguientes fases:

— Delimitación de los objetivos y escenarios.

— Identificación de las fuentes de datos, los procedimientos de acceso a los escenarios y las unidades de observación.

— Determinación de los procedimientos, instrumentos de recogida de información y los tiempos. Selección de informantes clave.

— Acceso al campo, a los escenarios. Recogida de información. Sistemas de registro.

— Análisis de la información. Sistemas de codificación y de representación de los datos.

— Sistemas para la validación de la información. La triangulación.

— Interpretación de los resultados, realización de inferencias, verificación.

— Elaboración del informe.

El orden no tiene por que ser fijo; los objetivos pueden verse ampliados y modificados; los modos de recogida de información pueden igualmente ampliarse; los procedimientos de análisis de datos se dan de modo simultáneo a su recogida, y la interpretación de los resultados y los sistemas de validación pueden aconsejar una ampliación de la recogida de información. De este modo, es posible que la investigación difiera del diseño original, influida por la propia realidad, por el punto de vista de los actores sociales que pueden modificar la perspectiva del observador, especialmente en las primeras etapas de la permanencia en el campo, lo que le permitirá definir el problema y los nuevos elementos para continuar el estudio. Y es que el contexto confiere un significado específico a los problemas que se estudian, dado que los hechos son significativos

desde el contexto en que ocurren, desde las personas que integran los grupos, desde sus relaciones y desde los códigos culturales de la propia comunidad para interpretarlos. La observación suele comenzar con un necesario conocimiento amplio, descriptivo, de lo que acontece, para poco a poco ir centrándose en los aspectos clave de la problemática a estudiar.

En la observación participante el propio observador es el principal instrumento, sus ojos y sus oídos son instrumentos esenciales, así como el diario o cuaderno de campo, donde se debe indicar el día, hora, lugar y persona o personas de los hechos, debiendo escribir las notas de campo inmediatamente ocurrido un suceso, en presencia de los observados, siendo importante anotar aquellos comportamientos no verbales que sirvan para confirmar lo que se ha dicho verbalmente. Otros instrumentos a utilizar son los documentos escritos, cuestionarios, cámaras fotográficas y de video, así como las entrevistas a informadores clave, historias personales y profesionales y encuestas.

3.5.3. El proceso de recogida de datos en la observación participante

La observación participante es una estrategia de campo amplia que comienza antes de introducirse en el campo, con la delimitación de la situación de lo que se pretende estudiar, los objetivos del estudio y el plan de lo que se va a hacer, el medio físico, las personas y sus interacciones, la concreción de los aspectos prácticos de inmersión en el campo y el contacto con los grupos. La observación combina, simultáneamente, el uso de varias técnicas de recogida y análisis de la información, como son el análisis de documentos, la entrevista a personas en la situación e informantes específicos, la participación directa, la observación y la introspección. El conocimiento de la historia previa a los acontecimientos que van a ser observados es fundamental para comprender el medio social y lo que en él acontece.

Observar lo que ocurre en una situación de estudio dada, las actividades, lo que dice la gente, cómo interactúa, la naturaleza del medio físico, son importantes en un enfoque comprensivo de trabajo de campo, pero también es importante observar lo que no ocurre y sería esperable y de-

seable que ocurriera; si, por ejemplo, se está estudiando un programa innovador y las metas del programa, su diseño de implementación y las propuestas de ciertas cosas que deben ocurrir o se espera que ocurran, no acontecen, debe ser anotado por el observador. Todo el proceso de observación estará influido por la preparación, sentido común y experiencia del observador.

Una de las funciones del observador consiste en realizar, mediante notas de campo, descripciones de lo que está siendo vivido y observado, y serán importantes para determinar la calidad de los análisis e interpretaciones posteriores. Las notas de campo son la base fundamental de los estudios de caso en la investigación cualitativa, y deben recoger citas literales de la gente observada, de sus conversaciones, de los sentimientos y reacciones del observador sobre lo que está observando y realizar síntesis y los «*insights*» o compresiones intuitivas que genera la observación y su interpretación (Figura 4). El observador debe seguir ciertas estrategias para hacer las anotaciones, las grabaciones, de modo que sea capaz de conseguir que su trabajo no afecte a su participación; una norma básica a tener en cuenta es escribir rápida y brevemente y completar las notas de campo tan pronto como sea posible fuera de la situación de campo. La extensión o la amplitud depende de las decisiones del observador. Las clases de registro que suelen utilizarse para organizar la información recogida son: el registro básico de notas tomadas durante las observaciones, entrevistas o grabaciones, que deben conservarse sin ninguna alteración; el registro temático, basado en las descripciones y reflexiones del investigador en torno a diversos campos temáticos; el registro logístico, reflejando las dificultades encontradas, los acontecimientos significativos, las reflexiones personales y los registros complementarios (López-Barajas, 1998:27).

Veamos un ejemplo de hoja de registro en la observación participante:

HOJA DE REGISTRO EN LA OBSERVACIÓN PARTICIPANTE

Lugar de la observación: ..... Fecha:.....

Nombre de observador: .....

— ¿Quiénes están en la escena?

— ¿Qué está sucediendo?

— ¿Qué interacciones se producen?

— ¿Qué estatus y roles aparecen en la interacción?

— ¿Cuál es el contenido de sus conversaciones?

— ¿Qué lenguajes verbales y no verbales utilizan para comunicarse?

— ¿Dónde está situado el grupo o la escena?

— ¿Con qué recursos cuentan?

— ¿Qué sensaciones y sentimientos se detectan en los contextos del grupo?

— ¿Qué significados atribuyen los participantes a su conducta?

— Otras informaciones ¿Cuál es la historia del grupo?

— Comentarios del observador.

Figura 4. Ejemplo de hoja de registro en la observación participante.

La observación participante combina la observación y la entrevista informal. El observador debe ser capaz de conocer el lenguaje específico de los ambientes que observa así como la comunicación no verbal. Durante las actividades informales no planificadas suele ser difícil organizar las observaciones; es imposible anticipar qué tipo de cosas pueden emerger, pero es importante recordar que todo lo que ocurre en torno a las cuestiones objeto de la observación puede ser de interés. Las entrevistas completan la información recogida mediante la observación, pero la información que aportan los entrevistados puede ser erróneas o ser autojustificativa, porque los participantes informan de sus percepciones y desde su punto de vista. Por eso las observaciones proporcionan una comprobación de lo que se recoge en las entrevistas; éstas permiten al

observador ir más allá del comportamiento externo para explorar los estados internos de las personas que han sido observadas.

Otra fuente de información son los documentos y registros sobre los temas de estudio, siendo importante que al comienzo del trabajo el investigador negocie su acceso a dichos documentos y archivos, dado que proporcionan información valiosa a la vez que estímulos para generar cuestiones a comprobar a través de la observación directa y de las entrevistas. Los análisis de documentos proporcionan otra visión de las escenas y su información puede ser utilizada para guiar y comprobar la información recogida por otras fuentes. Cada fuente de datos tiene sus ventajas y desventajas pudiendo un enfoque compensar la debilidad del otro.

El trabajo de campo no es un método o técnica simple. El observador está en el sitio en el que se desarrolla el programa observando y participando con la gente, entrevistándola y consultando los documentos básicos del programa. Y debe permanecer atento a todas las posibles fuentes de error, verificando la información recogida. La combinación de las técnicas de observación, entrevista y análisis de documentos, ayudará el trabajador de campo a aclarar los constructos desarrollados y a corregir los sesgos que aparecen cuando el fenómeno se estudia por un solo procedimiento. Todo junto resulta en un enfoque multimétodo, que contrastado mediante la triangulación incrementa la validez y la fiabilidad de los datos de la investigación.

3.5.4. El análisis de la información en la observación participante

Cuando el observador ha acumulado suficiente información para realizar el análisis, los procedimientos básicos son la codificación, la búsqueda de categorías de significado, la relación entre categorías, la interpretación de los datos y la exposición de los hallazgos utilizando tablas, diagramas, análisis estadísticos si fueran necesarios, intercalados en la redacción escrita que serán incluidas en el informe. Mediante la codificación de las informaciones recogidas se va reduciendo la información y haciéndola manejable para la búsqueda de categorías de significado, necesarias para su interpretación.

El análisis de los datos puede seguir los siguientes pasos:

1.º Lectura de la información textual.

2.º Asignación de códigos a las unidades de significado, indicándolos al margen del texto.

3.º Desarrollo de categorías para agrupar los códigos que representan segmentos e ideas con significado similar.

4.º Agrupación de categorías formando categorías de orden superior.

5.º Revisión de las categorías, creación de nuevas categorías y reagrupaciones y reasignaciones si fuera necesario.

6.º Relación entre las categorías, explicación según su procedencia original y su relación con otras teniendo en cuenta las situaciones y sus momentos.

7.º Ampliación de la información si fuera necesario para una comprobación de los datos y verificación de la información.

La interpretación de los resultados puede recorrer también diversos procesos según la perspectiva formal del análisis: de categorías, temático, taxonómico, componencial, hermenéutico, semiótico, temático, simbólico, de valores, fenomenológico, filosófico e histórico. Los datos pueden ser analizados utilizando los programas informáticos de análisis textual que existen en la actualidad, como son el Atlas ti o el SPSS, entre otros.

3.5.5. La elaboración del informe de la observación participante

El observador participante debe ser descriptivo en las notas de campo, reunir una variedad de información desde diferentes perspectivas, seleccionar los informantes clave para aprovechar sus conocimientos, sus habilidades, su posición, que pueden resultar de gran ayuda, usar citas representativas de los participantes del programa en sus propios términos, captando el punto de vista de los participantes, de sus experiencias, con sus propias palabras.

El observador, para la elaboración del informe, debe tener en cuenta los diferentes estadios del trabajo de campo, y en consecuencia deberá reflejar:

— Los motivos que le llevaron a interesarse por el tema y a iniciar el estudio.

— La revisión de la información disponible y la literatura relevante al tema.

— La formulación del problema a investigar y de los objetivos del estudio.

— Las negociaciones para estudiar la situación en el contexto.

— El plan de acceso al campo, el modo de introducirse en las situaciones y relacionarse con las personas y los grupos.

— Las técnicas de recogida de información utilizadas y los análisis de información y ampliación de la permanencia en el campo y la búsqueda de información complementaria.

— La validación de los hallazgos y de sus interpretaciones y devolución de las mismas al grupo.

— La presentación de los resultados, elaboración de las conclusiones y recomendaciones.

— La presentación del informe final.

Patton, (1990:273) considera que el observador participante debe tener en cuenta las siguientes recomendaciones:

— Comunicar la verdad e informar de un modo sincero en el estadio inicial al entrar a los sitios del trabajo de campo, teniendo en cuenta que el investigador-observador es también observado y evaluado.

— Ser disciplinado y consciente en la toma de notas detalladas en todos los estadios del trabajo de campo y permanecer alerta durante la mayoría las situaciones rutinarias de la fase intermedia del trabajo de campo.

— Estar tan implicado como sea posible en experimentar el programa de un modo global, mientras se mantiene una perspectiva analítica en el trabajo de campo.

— Centrarse en unificar la información, dentro de unas líneas determinadas, en una síntesis útil del trabajo de campo, separando, nítidamente, las descripciones de las interpretaciones y de los juicios.

— Proporcionar un feedback formativo como parte del proceso de verificación del trabajo de campo.

— Incluir en las notas de campo y en los informes las experiencias propias del observador, sus pensamientos y sus sentimientos. Éstos son también datos de campo.

— Cuidar la entrega de la información a tiempo y observar su impacto.

4. LA EVALUACIÓN DE NECESIDADES

Los estudios de evaluación de necesidades consisten en reunir información y analizarla con la finalidad de identificar las necesidades de los individuos, los grupos, las instituciones, las comunidades y la sociedad, e identificar las áreas en las que existen deficiencias, con el objetivo último de la utilización de dicha información para la realización de acciones, concretadas, en muchos casos, en programas de intervención tendentes a subsanar y mejorar la situación detectada, planificar su realización y valorar sus resultados para comprobar si se ha logrado satisfacer las necesidades y si es necesario iniciar nuevas mejoras. La valoración de necesidades se suele utilizar para desarrollar planes de acción inmediata o a medio y largo plazo.

La definición más utilizada de «necesidad» en el contexto de valoración de necesidades es la que se refiere a la discrepancia entre el nivel deseable y el conseguido, y se dice que existe una necesidad en un determinado campo cuando el nivel actual es inferior al deseado. Una segunda acepción de la definición de necesidad hace referencia al deseo de una preferencia. Y una tercera acepción de necesidad se basa en la deficiencia. Entonces, existe una necesidad cuando: 1) lo conseguido es inferior a

lo que se pretendía, 2) se desea algo que no se tiene y 3) se da una deficiencia que hay que subsanar (Kaufman, 1972).

La valoración de necesidades es considerada un componente de la evaluación, siendo una de las actividades de todo proceso de investigación evaluativa, circunscrita a la fase inicial y de la fase última conclusiva del mismo proceso evaluativo que cierra y abre un nuevo ciclo con la enumeración de áreas de necesidad prioritaria a tener en cuenta en los planes de mejora, dentro de los procesos de mejora continua de las instituciones y de los programas.

Los estadios o fases seguidos en los estudios de valoración de necesidades, son los siguientes:

1. *Afirmación de los propósitos de una valoración de necesidades:* Como otras formas de investigación, planificar un estudio de valoración de necesidades requiere clarificar los propósitos de valoración y delimitar las áreas en las cuales los estudios de valoración de necesidades se van a concentrar.

2. *Normas para la identificación de necesidades:* Cuando se planifica un estudio de valoración de necesidades en educación, las metas educativas, los resultados a lograr, son, frecuentemente, las normas frente a las cuales se compara el nivel obtenido para determinar las discrepancias. Otro modo de identificar las necesidades es la valoración subjetiva por parte de los informantes.

3. *Diseños de valoración de necesidades:* Los diseños para la valoración de necesidades comienzan con una especificación del objetivo del estudio, lo cual supone una delimitación de los propósitos específicos y los ámbitos en los que se van a valorar las necesidades. Para ello se requerirá elaborar instrumentos con los indicadores representativos de los propósitos y ámbitos específicos, recoger la información, analizarla y realizar informes. Los instrumentos para recoger opiniones suelen ser los propios de la de investigación tipo encuesta: cuestionarios, inventarios, hojas de registro, entrevistas.

4. *Asignación de prioridades de las necesidades:* El resultado de los estudios de valoración de necesidades es la identificación de muchas necesidades. Para maximizar la utilidad del estudio, las necesidades identi-

ficadas deben ser ordenadas de acuerdo con el grado de consenso alcanzado, según la importancia otorgada, la posibilidad que tiene de ser resuelta con éxito y el impacto que va a tener en la satisfacción de la necesidad.

5. *Uso de los resultados:* Una de las características de los estudios de valoración de necesidades radica en su finalidad que es la utilización de los resultados en aquello para lo que se ha realizado el estudio. Supone el último estadio del proceso: ser activo en el uso de los hallazgos.

Dado que los análisis para la valoración de necesidades se realizan en diferentes situaciones, las estrategias utilizadas también varían. Suelen utilizarse técnicas como el Delphi, los grupos de discusión, las entrevistas en profundidad, la observación, e instrumentos como cuestionarios, inventarios y hojas de registro.

5. EL CUESTIONARIO

El cuestionario es una técnica de recogida de información sobre opiniones, actitudes, habilidades, aplicable a amplias muestras representativas de poblaciones definidas. Su finalidad es recoger información de modo sistemático, de acuerdo con los objetivos de la investigación, o evaluación, sobre las variables que intervienen en la investigación. El término 'cuestionario' también designa al instrumento concreto de recogida de información, y, junto con la entrevista es una de las técnicas más utilizadas en la investigación tipo encuesta.

Los cuestionarios se presentan de forma escrita, en papel o en formato electrónico a través de internet, para ser completados por los propios encuestados, debiendo incluir instrucciones claras sobre cómo responderlos.

El criterio fundamental que guía la construcción de un cuestionario es el propósito de la investigación, sus objetivos, debiendo tener en cuenta la población a la que se destina, debiendo cumplir con las exigencias de los instrumentos de medida como son su fiabilidad y su validez.

Las principales fases para la elaboración de los cuestionarios son las siguientes:

1.º Delimitación del rasgo o rasgos a medir, de los contenidos, en coherencia con los objetivos de la investigación, concretado en las variables que se van a estudiar.

2.º Creación de los ítems. Enunciación de proposiciones o elementos iniciales representativos de las cuestiones que se quieren estudiar. Cada ítem debe centrarse en sólo un aspecto, de un modo claro, preciso, no ambiguo y no excesivamente largo.

Las preguntas que integran un cuestionario necesariamente tendrán que ver con los objetivos de la investigación, es decir, con la naturaleza de la información que se quiera recoger. Se debe evitar formular preguntas generales si lo que se pretende es recoger información específica. A veces es necesario formular diversas preguntas sobre una misma cuestión para alcanzar mayor precisión en la respuesta. Se deben incluir preguntas sobre hechos, acciones, de información y preguntas de intención.

Los tipos más frecuentes de preguntas que se pueden introducir en un cuestionario se clasifican en abiertas y cerradas. Las preguntas abiertas, también llamadas de *respuesta libre* se caracterizan por que al sujeto no se le imponen ninguna limitación sobre la categoría de respuesta, ni de espacio para responder. Las preguntas abiertas se suelen utilizar en investigaciones exploratorias en las que el investigador necesita conocer de un modo amplio la información que un determinado grupo posee sobre aquello que le interesa conocer y, posteriormente, buscará las categorías esenciales en torno al tema que investiga para, a partir de esta categorización, realizar un análisis de la información aportada. Las preguntas cerradas pretenden recoger el grado de acuerdo o desacuerdo del encuestado con la pregunta o afirmación formulada, pudiendo optar por diversos tipos de respuestas: de sí o no, de escala tipo Likert, con gradaciones, por ejemplo, de 1, mínimo acuerdo, a 5, máximo acuerdo con el ítem presentado.

3.º Análisis de los ítems por un grupo de jueces o expertos en la temática que se estudia. Valoración de las proposiciones por un grupo de jueces conocedores del tema sobre el que queremos recoger información,

que valorará la adecuación de las preguntas, o ítems, con el rasgo que pretendemos medir, con los objetivos, la adecuación en la formulación de los elementos a la muestra a la cual se dirige, su grado de inteligibilidad y la longitud o número de ítems de la prueba.

Una vez que la prueba, el cuestionario, ha sido valorada por el grupo de jueces, se aplica a un grupo piloto para el análisis de fiabilidad, validez, homogeneidad y dificultad de los ítems a partir de las respuestas dadas por los encuestados.

La fiabilidad, en este tipo de instrumentos, tiene que ver con la consistencia con que los elementos son contestados por el grupo, y a lo largo del tiempo. Para conocer la consistencia interna de la escala respecto a la actitud que se quiere medir se utiliza el coeficiente *alfa* de Cronbach.

La validez suele referirse a la de contenido, en clara referencia a la congruencia entre el ítem y el objetivo sobre el que se persigue recoger información. Para hallar la validez de la escala se puede utilizar el criterio de jueces así como un criterio externo para compararla, pudiendo ser otra escala que mida la misma actitud y haya sido ya validada.

Si se obtienen índices de fiabilidad y validez aceptables, el cuestionario estará listo para ser aplicado en la investigación, y sus datos serán adecuados para realizar el análisis de la información, valorar los resultados y elaborar conclusiones y recomendaciones.

4.º Elaboración del cuestionario definitivo. El cuestionario es, sin duda, el instrumento más utilizado en la investigación en ciencias sociales y en la evaluación. Puede adoptar muchas formas, por ejemplo, listas de control, escalas, preguntas de respuesta múltiple, preguntas abiertas. La mayoría de los cuestionarios incluyen una combinación de varios tipos diferentes de preguntas. También se puede dar una serie de respuestas posibles y pedir a los encuestados que las enumere en orden de preferencia.

Un cuestionario bien elaborado proporciona gran riqueza de información, siempre que se centre en recoger la información que se necesita, evitando que las personas que lo contesten sientan que han dedicado en

responder más tiempo del necesario. Mostramos algunos ejemplos de cuestionarios (Figuras 5 y 6).

Ejemplo de cuestionario sobre una sesión de clase:

CUESTIONARIO SOBRE LA SESIÓN DE CLASE

Este cuestionario tiene como objetivo conocer sus impresiones sobre la clase que acaba de finalizar. Los resultados se utilizarán con fines de mejora. Contestar al cuestionario le llevará alrededor de 10 minutos. Muchas gracias.

Exprese su grado de acuerdo sobre las proposiciones que se le muestran a continuación, de modo que refleje sus percepciones sobre la sesión de clase usando la siguiente escala:

1 = Muy en desacuerdo
2 = Muy en desacuerdo
3 = Ni de acuerdo ni en desacuerdo
4 = De acuerdo
5 = Muy de acuerdo

I. Actividades del profesor en el aula	
1. El profesor presenta al inicio de la clase los objetivos de la lección.	1 2 3 4 5
2. El profesor, a medida que explicaba los contenidos muestra ejemplos.	1 2 3 4 5
3. El profesor pregunta a algún alumno para comprobar su comprensión de la explicación.	1 2 3 4 5
4. El profesor muestra ejemplos de aplicación práctica de lo explicado.	1 2 3 4 5
II. Actividades del alumno en el aula	
5. Los alumnos preguntan sus dudas al profesor.	1 2 3 4 5
6. La interactividad fue esencial en el curso.	1 2 3 4 5
7. Los alumnos realizan las actividades propuestas por el profesor en el aula.	1 2 3 4 5

Figura 5. Ejemplo de cuestionario sobre una sesión de clase.

Cuestionario para la evaluación de un curso en línea:

CUESTIONARIO PARA LA EVALUACIÓN DE UN CURSO EN LÍNEA

Nombre del curso o módulo que evalúa: .....

Nombre del participante (o código personal):..... Fecha:.....

Instrucciones

Por favor indique su valoración en cada una de las cuestiones del cuestionario, teniendo en cuenta que el 1 representa el valor más bajo y la impresión más negativa, el 3 representa una impresión adecuada y el 5 representan el valor más alto y la impresión más positiva. Conteste a las preguntas abiertas para comentarios adicionales. Contestar el cuestionario no le llevará mas de 15 minutos Sus aportaciones serán tenidas en cuenta para la mejora del curso. Muchas gracias.

CONTENIDOS DEL CURSO	
1. Conocía los requisitos (conocimientos y habilidades) exigidos para realizar este curso.	1 2 3 4 5
2. Dominaba los conocimientos y habilidades requeridos para este curso.	1 2 3 4 5
3. Los objetivos del curso están claramente formulados.	1 2 3 4 5
4. El curso ha estado a la altura de mis expectativas.	1 2 3 4 5
5. Los contenidos han sido relevantes para mi trabajo.	1 2 3 4 5
LA REALIZACIÓN DEL CURSO	
6. Las actividades del curso estimularon mi aprendizaje.	1 2 3 4 5
7. La interactividad fue esencial en el curso.	1 2 3 4 5
8. Las actividades en el curso me proporcionaron suficientes prácticas.	1 2 3 4 5
9. El curso me proporcionó suficientes oportunidades para comprobar mis avances.	1 2 3 4 5
10. La evaluación del curso ha sido correcta.	1 2 3 4 5
11. El nivel de dificultad del curso es apropiado.	1 2 3 4 5
12. El ritmo de avance del curso es apropiado.	1 2 3 4 5
EL PROFESOR DEL CURSO	
13. El profesor está bien preparado.	1 2 3 4 5
14. El profesor estaba cuando se le necesitaba.	1 2 3 4 5
EL MEDIO DE APRENDIZAJE	
15. La facilidad de formación por este medio fue cómoda.	1 2 3 4 5
16. Fue fácil conseguir por este medio todo lo que necesitaba para aprender.	1 2 3 4 5
RESULTADOS DEL CURSO	
17. Conseguí los objetivos del curso.	1 2 3 4 5
18. Podré utilizar lo que he aprendido en el curso.	1 2 3 4 5
19. Lo mejor del curso y por qué:	
20. Lo peor del curso y por qué:	

Figura 6. Ejemplo de cuestionario para la evaluación de un curso en línea.

Cuestionario para la revisión de procesos. El «Registro para la revisión del proceso», o de evaluación formativa, es un instrumento sencillo que puede ser utilizado por cualquier persona relacionada, ya sea usuario, ya evaluador, con el objeto a evaluar, programa, curso, módulo, para ver como funciona durante el proceso. La utilización de un instrumento de este tipo por muchos y diferentes usuarios para valorar un curso durante su proceso de realización, proporcionará información valiosa para la mejora del curso. Mostramos un ejemplo de hoja de «Revisión procesual» (Figura 7), basada en Lewis (1995).

EVALUACIÓN DE LA USABILIDAD DE UN CURSO POR INTERNET

Nombre del curso: .....

Nombre del evaluador: ..... Fecha:.....

Instrucciones

Por favor, indique su grado de acuerdo con las afirmaciones que le presentamos, teniendo en cuenta que la escala va desde el 1, máximo desacuerdo, hasta el 7, máximo acuerdo. Marque NA, si no se aplica

CONTENIDOS DEL CURSO								
1. En general, estoy satisfecho con lo fácil que es usar el sistema.	1	2	3	4	5	6	7	NA
2. Pude realizar el curso eficazmente utilizando el sistema.	1	2	3	4	5	6	7	NA
3. Fui capaz de realizar las actividades a través del sistema.	1	2	3	4	5	6	7	NA
4. La información proporcionada por el sistema, es clara.	1	2	3	4	5	6	7	NA
5. El sistema de mensajes de error dice como debo arreglar los problemas.	1	2	3	4	5	6	7	NA
6. La información proporcionada por el curso fue fácil de comprender.	1	2	3	4	5	6	7	NA
7. Es fácil encontrar la información que necesité.	1	2	3	4	5	6	7	NA
8. La información me ayudó a completar las tareas.	1	2	3	4	5	6	7	NA
9. La organización de la información es clara.	1	2	3	4	5	6	7	NA
10. La interfaz del sistema es agradable.	1	2	3	4	5	6	7	NA
11. En conjunto, estoy satisfecho con el sistema.	1	2	3	4	5	6	7	NA
Enumere tres aspectos negativos del sistema: 1. 2. 3.								
Enumere tres aspectos positivos del sistema: 1. 2. 3.								

Figura 7. Ejemplo de cuestionario de evaluación de usabilidad de un curso por internet.

6. LA ENTREVISTA

La entrevista es otro de los procedimientos y técnicas utilizadas para la recogida de información utilizada en la investigación en Educación, las Ciencias Sociales y en la evaluación de programas e instituciones. Se diferencia del cuestionario en que suele realizarse de modo individual y oral, con la presencia del entrevistador y entrevistado o entrevistados. Pertenece a la investigación tipo encuesta, debiendo realizarse de un modo sistemático.

La característica esencial de la entrevista, y su aportación principal como técnica de recogida de información y datos, es que permite profundizar más allá de la formulación de la pregunta, e incluso añadir nuevas preguntas para obtener una información más completa y precisa, y esto hace de la entrevista una técnica compleja, requiriendo de una gran profesionalidad del entrevistador para que sea provechosa. Puede utilizarse para completar la información obtenida mediante otras técnicas como puede ser la observación, el cuestionario o la discusión grupal.

Las entrevistas tienen un grado diferente de estructuración en función de sus propósitos. Las entrevistas estructuradas son encuestas en las que el investigador entra en contacto personal con los encuestados y les hace una serie de preguntas previamente establecidas. Este tipo de entrevistas suelen realizarse siguiendo un guión estructurado en torno a la información que se pretende obtener, de acuerdo con los objetivos de la investigación, para evitar influir en la dirección de la respuesta. Las instrucciones a los entrevistadores deben estar contenidas en el protocolo de la entrevista. Así, por ejemplo, si se trata de una entrevista profesional, el guión se atendrá al perfil y competencias requeridas para el puesto de trabajo que se desea cubrir, y las preguntas guardarían estrecha correspondencia con el perfil del puesto de trabajo. En otros casos, las entrevistas pueden caracterizarse por no seguir un guión muy estructurado, centrándose más en el propio desarrollo de las respuestas dadas por el entrevistado. Si se tratara de una entrevista de investigación, la estructura y estilo de las preguntas se realizarían de acuerdo con los objetivos de la propia investigación.

Las entrevistas necesitan ser planeadas cuidadosamente para conseguir el tipo y calidad de información que se persigue. Mostramos un

ejemplo breve de «Protocolo de entrevista» (Figura 8), mediante una lista de preguntas en relación con un curso multimedia interactivo.

PROTOCOLO DE ENTREVISTA

Nombre del entrevistado: .....

Entrevistador:..... Lugar:..... Fecha: .....

1. ¿Cuál es su especialidad?
2. ¿Cuántos años tiene de experiencia en la profesión?
3. ¿Cuántos años tiene de experiencia en esta compañía?
4. ¿Describa como utiliza el curso multimedia?
5. ¿Describa sus primeras reacciones ante el curso?
6. ¿Mencione sus opiniones actuales sobre el curso?
7. ¿Necesita entrenamiento adicional para el aprovechamiento del curso?
8. ¿En qué grado ha conseguido usted los objetivos establecidos para el curso?
9. ¿Qué le diría a otra persona que estuviera interesada en el curso?
10. ¿Qué tipos de éxitos ha experimentado tras recibir el curso?
11. ¿Qué tipos de problemas ha experimentado en relación con el curso?
12. Describa las áreas en que usted se siente competente tras la relación del curso.
13. ¿Qué mejoras recomendaría en el curso?
14. ¿Cuál es que su opinión del sistema multimedia interactivo utilizado en este curso?
15. ¿Qué añadiría en el curso?
16. ¿Qué quitaría en el curso?
17. ¿Algo más que decir sobre el curso?

Figura 8. Ejemplo de protocolo de entrevista.

7. EL ANÁLISIS DE CONTENIDO

El análisis de contenido es una técnica de investigación destinada a formular, a partir de ciertos datos, normalmente recogidos en un soporte físico como es un documento escrito, una grabación, u otro medio, inferencias fiables y válidas aplicables a su contexto. El análisis de contenido

ha llegado a ser un método científico capaz de ofrecer inferencias a partir de datos esencialmente verbales, simbólicos o comunicativos.

7.1. Concepto y alcance del análisis de contenido

Para Ander-Egg el análisis de contenido es una técnica de recopilación de información que permite estudiar el contenido manifiesto de una comunicación, clasificando sus diferentes partes, conforme a las categorías establecidas por el investigador, con el fin de identificar de manera sistemática y objetiva dichas categorías dentro del mensaje (1987:330).

Intuitivamente, el análisis de contenido podría caracterizarse como un método de investigación del significado simbólico de los mensajes. Sin embargo, hay que decir que los mensajes no siempre tienen un único significado, pudiendo contemplarse desde múltiples perspectivas y, además, no es necesario que exista coincidencia acerca de los significados, si bien el consenso o el acuerdo intersubjetivo sobre lo que significa un mensaje simplifica el análisis de su contenido, pero difícilmente la coincidencia puede servir como presupuesto para el análisis de contenido, puesto que esta coincidencia depende de si se comparte la misma perspectiva cultural o sociopolítica.

El rasgo más característico de los mensajes en el contenido es que informan a una persona de forma vicaria, proporcionando al receptor un conocimiento acerca de hechos o sucesos que se producen en un lugar lejano acerca de objetos que quizá existieron en el pasado o acerca de las ideas de otras personas. La naturaleza vicaria de las comunicaciones simbólicas es lo que obliga al receptor a formular inferencias específicas a partir de los datos que analiza en relación con su medio empírico llamado contexto de los datos.

Las teorías del significado y de los fenómenos simbólicos se asemejan entre sí por la importancia que conceden a la relación entre los datos y su contexto. Los datos son estímulos físicos, pero el interés por su significado parte de un nivel superior de abstracción.

El análisis de contenido establece el objeto de investigación y sitúa al investigador en una posición concreta frente a la realidad. El marco refe-

rencial dentro del cual puede representarse el papel que desempeña el investigador tiene tres finalidades: prescriptiva, analítica y metodológica. Es prescriptivo, en el sentido de que debe guiar la conceptualización y el diseño de los análisis de contenido práctico en cualquier circunstancia; es analítico en el sentido de que debe facilitar el examen crítico de los resultados del análisis de contenido efectuado por otros; y es metodológico en el sentido de que debe orientar el desarrollo y perfeccionamiento sistemático de los métodos de análisis de contenido.

7.2. Las fases del análisis de contenido

La finalidad del análisis de contenido es proporcionar conocimientos, nuevas compresiones y guía practica para la acción en relación con el tema que estudia. Como todo proceso de investigación, debe iniciarse con la enunciación clara de la finalidad y objetivos de lo que el analista quiere conocer y conseguir, y dado que el análisis de contenido suministra un conocimiento vicario, es decir, una información acerca de algo que no puede observarse directamente, el objetivo debe quedar situado dentro de su contexto y de los datos disponibles (Krippendorff, 1990).

Para la realización del análisis de contenido se señalan las siguientes fases:

1. Establecer la unidad de análisis: un documento, un artículo, las narraciones sobre un tema específico aportadas por un grupo, las sesiones de clase, las anotaciones de la discusión en grupo, las anotaciones de la observación participante, una grabación en vídeo.
2. Seleccionar la muestra del material de análisis: las anotaciones de tres sesiones de clase, los cinco primeros minutos, intermedios y finales de una grabación de vídeo, etc.
3. Determinar las categorías de análisis: según el objetivo del análisis de contenido, de la temática, serán necesarias diversas categorías: de significado, de valor, de personas. Las categorías vendrán determinadas por el propio contenido, por sus actores, por la frecuencia de aparición de determinados significados.

4. Realizar inferencias a partir de las categorías y su relación entre ellas, elaborando teorías explicativas.
5. Elaborar el informe, explicando los objetivos pretendidos, los materiales, el contexto, el proceso, los resultados, las conclusiones, y las fuentes documentales para verificar las conclusiones. Si el objetivo del análisis de contenido está enunciado de manera clara, inequívoca, podrá juzgarse si se ha podido realizar, adjuntando las pruebas, las fuentes de datos, que permitirán validar los resultados.

Todo análisis de contenido debe realizarse en relación con el contexto de la información y los datos, y justificarse en función de éste.

7.3. Las unidades de análisis y su contexto

El elemento esencial de la comunicación es la palabra. La unidad de análisis en el análisis de contenido es el elemento concreto de comunicación que se va a analizar. Se pueden considerar dos tipos de unidades básicas: las de tipo gramatical y las no gramaticales. En cuanto a las primeras, se suelen analizar los vocablos, especialmente teniendo en cuenta su significado dentro de las frases, su frecuencia de aparición y la relación de significado de unos con otros, así como el modo en que viene expresado. Las segundas hacen referencia a otros modos de manifestar la comunicación como es a través de fotografías, gráficos o de documentos audiovisuales.

En todo análisis de contenido, debe quedar claro qué datos se analizan, de qué manera se definen y de qué población se extraen. Los datos son los elementos básicos, primitivos, del análisis de contenido, y constituyen la información que el analista debe tratar de comprender. Los datos pueden venir dados mediante documentos escritos, pudiendo ser comunicados, documentos de prensa, documentos oficiales, documentos personales, documentos gráficos y audiovisuales, obras literarias, cinematográficas, sobre los cuales el analista realiza un análisis a posteriori. Pero también puede realizar el análisis de la comunicación en un contexto real, en vivo, sobre la cual el investigador observa, toma notas y registra mediante instrumentos técnicos, combinando la observación y el análisis textual, completándose el análisis con la revisión de las notas y

la audición y visionado de las grabaciones. En la realización a esta segunda situación del análisis de contenido hay que tener en cuenta los siguientes elementos:

— El emisor, persona que comunica o los productores del contenido y las causas de la comunicación, identificando sus intenciones, y determinando las características de las personas y de los grupos que emiten el mensaje.

— El código, o lenguaje empleado para establecer la comunicación, el mensaje o contenido de la comunicación, lo que se quiere decir y el canal, o medio por el cual se emite la comunicación. Las características del contenido, como son describir tendencias en el contenido de la comunicación, comparar los distintos niveles de comunicación, el estudio de su significado, el estudio de las formas de estilo.

— El efecto del contenido, en relación con el receptor o persona o público al que se dirige el contenido de la comunicación, o que la recibe, reflejando sus intereses, actitudes y valores, los aspectos que más le llaman su atención del contenido y las actitudes y respuestas emitidas ante las comunicaciones recibidas.

— El contexto en el que se produce la comunicación debe hacerse explícito y descrito por el analista con el fin de incluir todas las circunstancias circundantes, antecedentes o consecuentes al acto de la comunicación.

Todo trabajo de investigación debe definir los límites más allá de los cuales no podrá extenderse el análisis.

7.4. La inferencia como objetivo intelectual del análisis de contenido y la validez como criterio de éxito

En todo análisis de contenido, el objetivo último es la formulación justificada de inferencias a partir de los datos originales, teniendo en cuenta el contexto. Para realizar y justificar esas inferencias, el analista de contenido debe contar con relaciones relativamente estables entre los datos y el contexto, o construir una teoría operacional de esas relaciones, incluida la aportación de los factores mediadores. Se realizará una cons-

trucción analítica teórica de esas relaciones, formuladas de manera que los datos aparezcan como sus variables independientes y el objetivo forme parte de sus variables dependientes. Una construcción analítica ofrece reglas de inferencia sirviendo como puente lógico entre los datos disponibles y la finalidad incierta situada en su contexto, debiendo conocer la forma específica de esas relaciones.

La codificación de unidades textuales mediante categorías conceptuales se realiza a partir de inferencias, siendo estas inferencias la principal preocupación del análisis de contenido, dado que es una técnica de investigación para formular inferencias identificando de manera sistemática y objetiva ciertas características especificadas dentro de un texto.

Más allá de su continuo compromiso con cuestiones educativas, psicológicas, sociológicas y políticas, el interés por el uso de la técnica del análisis de contenido viene unido a la importancia de establecer criterios adecuados de validez. En todo análisis de contenido hay que especificar con claridad el tipo de pruebas necesarias para validar sus resultados, las inferencias sobre los fenómenos que se estudian, con el fin de que otras personas puedan comprobar si las inferencias son exactas. La metodología del análisis de contenido sólo puede progresar si se efectúan esfuerzos sistemáticos para validar sus resultados. Los análisis de contenido que no ofrecen pruebas adecuadas de su validez, resultan empíricamente inútiles y contribuyen poco al desarrollo de la metodología del análisis de contenido.

8. LA TÉCNICA DELPHI

Es una técnica originalmente utilizada para recoger información de las aportaciones personales de un grupo de expertos sobre los diferentes aspectos que definen un determinado problema y el modo posible de solucionarlo, haciendo partícipes de las opiniones subjetivas, de una manera anónima, a los otros miembros, con el fin de obtener el mayor grado de consenso para solucionar dicho problema, estableciendo prioridades. Su nombre procede del 'Oráculo de Delfos'. Se utiliza para el análisis de problemas diversos relacionados con la planificación y toma de decisiones, muy utilizada en el mundo empresarial y también en el socio-políti-

co, para analizar tendencias y anticiparse a problemas futuros, ofreciendo soluciones y planificando el modo de conseguirlas.

Los resultados aportados sobre las necesidades en torno a una determinada cuestión, y las soluciones posibles aportadas, guiarán la elaboración de planes y programas específicos dirigidos a dar respuesta a dichas necesidades. Así, por ejemplo, si se trata de recoger las opiniones y valoraciones de expertos para tratar de abordar y solucionar un problema como puede ser la integración escolar de alumnos procedentes de diferentes países, las aportaciones a través de un cuestionario, y las opiniones aportadas en las cuestiones abiertas dadas por un grupo suficientemente amplio y procedente de diversos campos, como puede ser el formado por maestros, pedagogos, psicólogos, sociólogos, políticos, educadores sociales, médicos, padres, aportaran suficiente información para el análisis y propuestas de soluciones fundamentadas y contratadas entre diferentes expertos y grupos de interés. Como puede deducirse por el ejemplo, los expertos lo son en el sentido de que deben dar su opinión sobre la cuestión a estudio, bien por su formación sobre el tema, bien por su relación con él, pudiendo ser considerados como informadores clave.

Desde la perspectiva de la búsqueda del consenso, la técnica Delphi, o Delfos, es considerada como una técnica de recogida de información cualitativa, aunque utiliza el instrumento de recogida de información del cuestionario con preguntas formuladas de modo abierto, pero también cerrado, similar al utilizado en las escalas tipo Likert; para el análisis de las cuestiones abiertas utiliza el análisis de contenido y para las cerradas, los análisis estadísticos.

El proceso de elaboración y recogida de información y datos mediante el Delphi es el siguiente:

La técnica Delphi, enmarcada en un proceso de investigación, debe comenzar, como siempre, con la formulación precisa del problema a estudiar, para continuar con una cuidada planificación de los pasos a realizar para resolverlo. El muestreo de expertos es así mismo, esencial, y los criterios para su selección deben ser especificados y argumentados. El modo de dirigirnos a ellos, el tipo de instrumento y los procedimientos para el análisis de la información y devolución de la información,

son los pasos siguientes, y también deben estar previstos en la fase de planificación del trabajo (Padilla Carmona, Moreno Sánchez y Vélez González, 2002).

Una vez concretado el tema de estudio, y profundizado en él a través de la revisión del estado de la cuestión, a continuación hay que pensar en los expertos que por su posible formación, experiencia e implicación en el tema de estudio, pueden aportar su opinión, para estudiar con la mayor profundidad el problema y la búsqueda de soluciones. El grupo de expertos debe ser lo suficiente amplio, para garantizar la representatividad de sus opiniones, y en número suficiente, teniendo en cuenta la perdida de informantes que a lo largo del proceso de consulta pueda producirse.

El grupo de expertos, lo es, desde diversas perspectivas. Si se pretende que el grupo tenga cierta homogeneidad en su especialización, puede ser suficiente entre 10 y 15 miembros; si lo que se requiere es diversidad de expertos, por que así lo aconsejan las características del problema a estudiar, el número de expertos a consultar será mayor, entre 30 y 50, procurando que de cada especialidad haya tres o más, para enriquecer las aportaciones, y prever, asimismo, la pérdida de informantes durante el proceso.

Después deberemos dirigirnos a cada miembro seleccionado, indicándoles los motivos para pedirle su colaboración, y explicándoles con claridad y precisión nuestro proyecto de investigación, las características de la técnica Delphi, y lo que esperamos de él, garantizándole la discrecionalidad en el uso de sus aportaciones, opiniones y valoraciones, así como el tiempo que se prevé que dure el proceso de consulta. El modo de dirigirnos a los informantes puede ser a través de correo electrónico, o cualquier otro. Durante todo el proceso debemos cuidar la presentación de nuestros objetivos de la manera más motivadora posible para conseguir y mantener la colaboración de los informantes o expertos.

El instrumento para la recogida de información y datos que utiliza es el cuestionario individual, normalmente enviado en formato electrónico, por correo electrónico y, tras el análisis de las respuestas emitidas por el grupo de expertos, se les reenvían los resultados grupales, junto con nuevas cuestiones, con el objetivo de alcanzar una mayor profundidad en las

tentativas de solución del problema, a la vez que el mayor grado de consenso posible. El reenvío de las opiniones y soluciones aportadas por los expertos puede hacerse hasta dos o más veces. Los expertos consultados no tienen por qué conocerse entre sí; la confidencialidad en la consulta debe garantizarse. Aunque no hay interacción personal entre los miembros del grupo, si es posible la influencia mutua dado que se comparte la información de las aportaciones del grupo.

Una vez aceptada la colaboración se enviará el primer cuestionario, con la información de los plazos para su devolución. El primer cuestionario se recomienda que siga un formato de preguntas lo suficientemente abierto para que se pueda expresar el mayor número de ideas posibles, desde su propia perspectiva. Sobre los cuestionarios contestados se realizará el análisis de la información, buscando el consenso en las aportaciones así como las ideas únicas y originales. El análisis de las preguntas abiertas puede realizarse mediante el análisis cualitativo de datos textuales, codificando la información y buscando las categorías de significado común a las aportaciones, y las ideas originales. Con ello se elaborará un segundo cuestionario, con un mayor número de preguntas cerradas, formadas a partir de los ítems sugeridos por las categorías del análisis del primer cuestionario, que será enviado a los expertos para recoger su valoración.

El segundo cuestionario estará formado por ítems redactados de un modo preciso, para que el experto emita sus respuestas mostrando su grado de acuerdo mediante escalas tipo Likert, u otras en las que se ordenen de acuerdo a su importancia en la incidencia de la cuestión de estudio, dejando posibilidades para que el experto añada cuestiones y sugerencias.

El análisis de los datos del segundo cuestionario se realizará utilizando las técnicas adecuadas al tipo de preguntas, y el posible nivel de medida que alcance. Así podrán utilizarse análisis estadísticos basados en el nivel de medida ordinal, o de intervalo para ofrecer puntuaciones de tendencia central, de dispersión, o incluso contrastes de grupos, junto con representaciones gráficas, en la búsqueda de acuerdos. En los resultados, se mostrarán los acuerdos logrados, junto con las propuestas divergentes argumentadas, que se enviarán como tercer cuestionario. Dado que con la técnica Delphi se pretende la búsqueda de consenso, la medi-

da más adecuada como indicador de consenso sería la moda, como la expresión de la opinión que más se repite, más que la media aritmética, pudiendo utilizarse tablas de porcentajes.

Los resultados aportados por los expertos, finalmente, serán publicados, y se utilizarán para la elaboración de las propuestas de mejora, basadas en muchos casos, en planes concretados en programas de intervención, y seguimiento de los mismos, para tratar de solucionar el problema que originó el estudio, y las necesidades detectadas.

Entre las ventajas de la técnica Delphi destacan la posibilidad de consultar a diversos expertos dispersos a lo largo y ancho de la geografía sin necesidad de reunirlos, la posibilidad de que cada uno de ellos manifieste sus puntos de vista sin influencia de los otros y la oportunidad de tras la primera consulta abierta y libre, tener la información de otros expertos para considerarla, contrastarla con la suya propia, y hacer síntesis enriquecedoras que favorezcan la amplitud y profundidad para atisbar los problemas, anticiparse a las necesidades y promover planes y programas de intervención para tratar de resolverlos.

9. LOS GRUPOS DE DISCUSIÓN

La técnica de los grupos de discusión se utiliza para la recogida de información de grupos de sujetos, creado de modo artificial, con la única finalidad de recoger información sobre un tema específico. Los miembros del grupo son reunidos durante un periodo de tiempo limitado, con el fin de debatir sobre un tema específico propuesto por el investigador. Los miembros del grupo deben ser homogéneos con el fin de posibilitar una discusión libre entre ellos, recomendándose, en algunos casos, que no se conozcan, para evitar prejuicios en su participación. Los datos de análisis serán los aportados en la discusión, por lo tanto, fundamentalmente textuales. Una vez profundizado, a lo largo de las reuniones que fueran necesarias, en el tema en cuestión, y obtenidos acuerdos y conclusiones, el grupo ha terminado con la misión para la que fue formado.

Como la técnica Delphi, el grupo de discusión suele utilizarse para profundizar en un tema, detectar las posibles necesidades sobre la cues-

tión objeto de estudio, de modo que se aporten posibles soluciones para elaborar planes, programas y acciones concretas de intervención. También puede utilizarse para evaluar programas ya en acción, y valorar sus resultados, efectos, y necesidades de mejora.

Las técnicas que venimos comentando en este tema son perfectamente complementarias entre sí, de modo que, cuando se está investigando una cuestión, un programa, puede utilizarse la observación participante, la consulta a un grupo de expertos mediante la técnica Delphi, la entrevista en profundidad a informantes clave, conocedores del tema, o la formación de un grupo de discusión formado por especialistas en el tema o cuestión, como pueden ser los profesionales que están aplicando un programa, un plan de estudios, o lo que se esté investigando o evaluando.

Las primeras fases del trabajo en el grupo de discusión suelen ser fundamentales para explorar de un modo amplio la problemática que se estudia, y a partir de ello ir focalizando, concretando, consensuando, los aspectos claves del tema, que guiarán la formulación de posibles hipótesis a comprobar, la realización de instrumentos de recogida de información como el cuestionario, guiones de entrevista o pautas de observación. También se utilizan los grupos de discusión para comentar los resultados obtenidos tras la aplicación de un programa piloto, para valorar los resultados obtenidos tras la aplicación de determinados instrumentos de recogida de información, para comentar el progreso en alguna investigación y para valorar los informes de resultados de un programa y las propuestas de mejora del propio programa y de aplicación.

El responsable de la investigación suele ser el moderador de la reunión, si bien se puede asignar un coordinador que puede hacer el papel de relator, tomando nota de los acuerdos y propuestas de trabajo a realizar para la siguiente reunión. El coordinador o el relator deberá tomar notas de lo que acontece en la discusión grupal, y realizar un informe de progreso que será tenido en cuenta para la redacción del informe final, el cual deberá reflejar el punto de partida, los objetivos, las hipótesis, el trabajo realizado en los distintos períodos, los instrumentos para la recogida de datos que se hayan elaborado y el análisis de la información.

Todo ello se incorporará en el informe general del trabajo de investigación o de la evaluación que se esté realizando.

Una vez que se ha profundizado, a lo largo de las reuniones necesarias, en el tema en cuestión, y obtenido el mayor grado de acuerdo, el grupo ha podido terminar con la misión para la que fue formado. Dado que los datos aportados en la discusión son, fundamentalmente, textuales, la técnica del análisis de contenido es la más adecuada para el análisis de la información aportada por el grupo.

Como técnica de recogida de información, los grupos de discusión, como la técnica Delphi, o la entrevista grupal, forma parte de las técnicas de encuesta. El Delphi difiere del grupo de discusión en que, para evitar las posibles influencias de unos miembros sobre otros, el investigador consulta a los seleccionados para que den sus opiniones y argumentos sin que unos y otros se conozcan, siendo el propio investigador, como coordinador, el que recoge las aportaciones y va devolviendo las síntesis y los aspectos más llamativos y controvertidos, en sucesivas vueltas de consulta, tantas como el coordinador considere necesarias hasta llegar a un cierto grado de comprensión o explicación sobre el tema que se estudia.

El enfoque de grupos de discusión necesita ser planificado cuidadosamente para conseguir el tipo de información y calidad que se persigue. Las preguntas deberán ser amplias, aunque contextualizadas en el tema que nos preocupa. Deberán ser lo suficientemente abiertas para que cada miembro del grupo pueda expresar su opinión. Y a lo largo del proceso se deberán ir concretando para aportar propuestas de acción para la solución del tema de estudio.

Para la formación y funcionamiento de los grupos de discusión, conviene tener en cuenta alguna de estas indicaciones:

— El tamaño del grupo, su composición, y posible número de grupos, en función de la temática y especificidad.

— Las característicos de los sujetos que deben formar parte de los grupos, contactar con ellos y conseguir su compromiso de colaboración y participación.

— Definir el papel de cada sujeto en su colaboración en la investigación y conseguir su compromiso.

— Delimitar la duración del proceso total de la investigación, así como el número de reuniones a mantener, el lugar y su duración.

Dado que se trata de recoger información amplia y especializada de los diferentes miembros que formen el grupo, y a la vez se busca homogeneidad, acuerdos, habrá que buscar tantos grupos como aspectos se consideren de interés para abordar el tema en profundidad y diversidad. Las reuniones en los grupos homogéneos se mantendrán hasta que se llegue al consenso, después algún representante de los grupos homogéneos se reunirán en el grupo multidisciplinar para avanzar desde cada perspectiva que integra la investigación, en la misma. El tipo de muestro es, por tanto, deliberado, o intencional, según el propósito de la investigación. El tamaño de los grupos suele aconsejarse entre 6 y 12. Las características personales y profesionales de los miembros de los grupos es de gran importancia, por que de ello depende el éxito en la investigación. Deben ser personas preparadas en el tema de estudio, abiertas al diálogo, responsables con sus compromisos. Estos compromisos son más fáciles de alcanzar, dado que la colaboración comporta un esfuerzo en tiempo y trabajo, si los miembros reciben algún tipo de incentivos, que cuando se trata de investigación subvencionadas, suelen ser incentivos económicos y de autoría en los hallazgos.

La duración de las reuniones suele ser de una hora a dos horas, como máximo, y debe estar fijada tanto la hora de inicio como de finalización. El responsable de la investigación suele ser el moderador de la reunión, si bien se puede asignar un coordinador para cada uno de los diferentes grupos, que puede hacer a la vez el papel de relator, que va tomando nota de los acuerdos, o estas notas pueden ir realizándose por otro miembro del grupo, o dependiendo de las características del mismo, ir consensuando y concretando los acuerdos y las propuestas de trabajo pendientes a realizar para ir avanzando en la investigación. También se puede realizar grabación de las discusiones por medios técnicos como son cámaras de video o magnetófono. Es conveniente mantener contactos durante los procesos intermedios a las reuniones de trabajo, pudiendo realizarse mediante acuerdos de colaboración y comunicación a través del correo electrónico o cualquier otro medio.

Al final, los representantes de cada uno de los grupos serán los responsables de elaborar los informes de progresos y avance, y contribuir a la redacción del informe final de la investigación. El informe deberá reflejar todo el proceso de la investigación, el punto de partida, los objetivos, las hipótesis, el trabajo de campo realizado en los períodos intermedios cuando se trata de un grupo de investigación, los instrumentos para la recogida de datos, el análisis de la información la y su incorporación en el informe general del trabajo de investigación.

Como puede verse se ha expuesto la técnica de los grupos de discusión para su utilización con diferentes fines. Para profundizar en un determinado tema, o como estrategia para la puesta en común procesual entre los miembros de un grupo de investigación.

EJERCICIOS DE AUTOEVALUACIÓN

1. Defina la observación sistemática recogiendo sus notas características.
2. Elabore un guión para la realización de la observación de una sesión de clase.
3. Realice una observación participante en el contexto en el que Vd. trabaje, utilizando, y completando, el modelo de registro de información que viene en el apartado 11.1.3, referido al proceso de recogida de datos en la observación participante.
4. Elabore un cuestionario para la evaluación de un programa educativo que conozca.
5. Características y finalidad del análisis de necesidades.
6. Características de los grupos de discusión, y valore su utilidad en los ámbitos de trabajo de la evaluación de programas.
7. Características de la técnica Delphi, y reflexione sobre las posibilidades de su aplicación en los ámbitos de trabajo del evaluador de programas.
8. Elabore un guión para realizar un seguimiento de una discusión grupal sobre un tema objeto de evaluación, como pueda ser la reunión de un

claustro de profesores para discutir un asunto de máxima importancia en un Centro Escolar como puede ser el acoso escolar. O la incorporación de las nuevas tecnologías para la enseñanza de las Ciencias. O cualquier otro tema o programa con el que esté familiarizado.

REFERENCIAS BIBLIOGRÁFICAS

Ander-Egg, E. (1987). *Técnicas de investigación social*. México: El Ateneo.

Caplow, T. (1972). *La investigación sociológica*. Barcelona: Laia, Ediciones de Bolsillo.

Evertson, C. M. y Green, J. L. (1986). «Observation as inquiry and method», en M. Wittrock (Ed.): Handbook of research on teaching. N. Y. London: MacMillan. (162-213)

— (1997). «La observación como investigación y método. En M. Wittrock, M. (Ed.): *La investigación de la enseñanza* I y II. Ed. Paidós-MEC. Madrid.

Goetz, J. P. y M. D. Lecompte (1988). *Etnografía y diseño cualitativo en investigación educativa*. Madrid: Morata, S. A.

Kaufman, R. A. (1972). *Educational system planning.* Englewood Cliff, N. J.: Prentice Hall.

Krippendorff, K. (1990). *Metodología de análisis de contenido.* Barcelona, Buenos Aires, Paidós.

Lewis, J. R. (1995). «IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use». *International Journal of Human-Computer Interaction*, 7, 1, 57-78.

López-Barajas Zayas, E. (Coor.) (1998): *La observación participante*. Madrid: UNED. Actas y Congresos.

Malinowski, B. (1922). *Argonauts of the western Pacific.* (Los Argonautas del Pacífico occidental). London: Routledge & Kegan Paul.

Martínez Mediano, C. (1996). Evaluación de programas educativos. Modelos de evaluación de programas. Investigación evaluativa. Madrid: *Cuadernos de la UNED.* 2.ª ed., 1997; 3.ª ed., 2007.

Martínez Mediano, C. (Coord.) (2004). Técnicas e instrumentos de recogida y análisis de datos. Madrid: Unidades Didácticas de la UNED.

Padilla Carmona, M. T., Moreno Sánchez, E. y Veléz González, E. (2002). «La técnica Delphi en la evaluación de necesidades: una aplicación al tratamiento del género en los centros escolares». *Bordón. Revista de Pedagogía*. 54, 83-94.

Park, R. y Burgess, E. (Eds.) (1921). *Introduction to the science of sociology.* Chicago: University of Chicago Press.

Patton, M. Q. (1990). Qualitative Evaluation and Research Methods. Newbury Park, CA.; Sage.

Postic, M. y De Ketele, J. M. (1992). *Observar las situaciones educativas*. Madrid, Narcea.

Ruiz Olabuénaga, J. I. (1999). *Metodología de la investigación cualitativa*. Bilbao: Universidad de Deusto.

Suárez Ortega, M. (2005). *El grupo de discusión. Una herramienta para la investigación cualitativa.* Barcelona. Alertes Educación.

Taylor, S. J. y Bogdan, R. (1998). *Introducción a los métodos cualitativos de investigación*. Barcelona: Paidós.

Tema 8

Técnicas para el análisis estadístico de la información en la evaluación de programas

1. Objetivos y resultados de aprendizaje.
2. El análisis de los datos en la investigación cualitativa.
   - 2.1. Descubrimiento en progreso.
   - 2.2. Codificación de los datos.
   - 2.3. Relativización de los datos.
3. Criterios para valorar las investigaciones cualitativas.
   - 3.1. Criterios de credibilidad, seguridad, confirmabilidad y transferibilidad.
   - 3.2. Criterios hermenéuticos.
   - 3.3. Criterios de autenticidad.
4. El análisis de datos en la investigación cuantitativa.
   - 4.1. Distribuciones numéricas y reducción de datos.
   - 4.2. Niveles de medida y tipos de variables.
   - 4.3. La distribución continua y la distribución normal de probabilidad.
   - 4.4. Representaciones gráficas.
   - 4.5. Medidas de tendencia central. La moda, la mediana y la media.
   - 4.6. Medidas de posición. Los cuartiles, deciles, percentiles.
   - 4.7. Medidas de variabilidad. La amplitud, la desviación media, la varianza y la desviación típica.
   - 4.8. Puntuaciones individuales. Las puntuaciones diferenciales y las puntuaciones típicas.
5. Las medidas correlacionales bivariadas y su interpretación.
   - 5.1. La correlación producto momento de Pearson.
   - 5.2. La correlación ordinal o de rangos de Spearman.

6. Comparación entre grupos. Los diseños de grupo.
   6.1. Concepto de experimento y de diseño de grupo experimental.
   6.2. Resolución de un diseño investigación de dos grupos independientes relacionados y nivel de medida de intervalo.
   6.3. Resolución de un diseño investigación de dos grupos relacionados y nivel de medida de intervalo.

Ejercicios de autoevaluación.

Referencias bibliográficas.

1. OBJETIVOS Y RESULTADOS DE APRENDIZAJE

Con este tema se pretende que el alumno comprenda la función que los procedimientos del análisis de datos cumplen en la investigación evaluativa y evaluación de programas, se espera que el alumno, tras el estudio de este tema, consiga los siguientes objetivos de aprendizaje:

— Revisar los contenidos relacionados con el análisis de datos, tanto cualitativos como cuantitativos aplicados a la evaluación de programas.

— Revisar los pasos para la realización de una investigación naturalista o cualitativa.

— Revisar el modo de reducción de los datos como primer acercamiento a su análisis.

— Revisar las diferentes representaciones gráficas.

— Revisar los estadísticos descriptivos de tendencia central, de posición, de variabilidad y dispersión y las puntuaciones individuales.

— Revisar el significado de diferentes coeficientes de correlación.

— Revisar los procedimientos seguidos en la resolución de los diseños de investigación de dos grupos.

2. EL ANÁLISIS DE LOS DATOS EN LA INVESTIGACIÓN CUALITATIVA

El enfoque cualitativo se orienta hacia el desarrollo de una comprensión profunda de los escenarios o personas que se estudian. El dato cua-

litativo nos ofrece una información básica, primaria, acerca de la realidad, con un bajo nivel de estructuración teórica. Estos datos no permiten, de una manera directa, establecer significados para resolver el problema de estudio. Junto con las notas de campo aparecen juicios, opiniones, reflexiones e interpretaciones del investigador haciendo que el análisis y la recogida de datos se alternen o superpongan en el proceso de investigación. La consideración de los procesos de recogida de datos y análisis de los mismos como fases separadas es poco productiva, defendiendo la concomitancia de dichas tareas en un proceso interactivo de recogida y análisis de datos que prosigue una vez finalizada la recogida de datos.

El análisis de datos cualitativos es una actividad que implica reflexiones, transformaciones, comprobaciones realizadas a partir de la información obtenida con el fin de obtener un significado relevante para el problema de investigación que se estudia. Miles y Huberman (1984: 23) representan los componentes del análisis de datos tal como se muestra en la Figura 1.

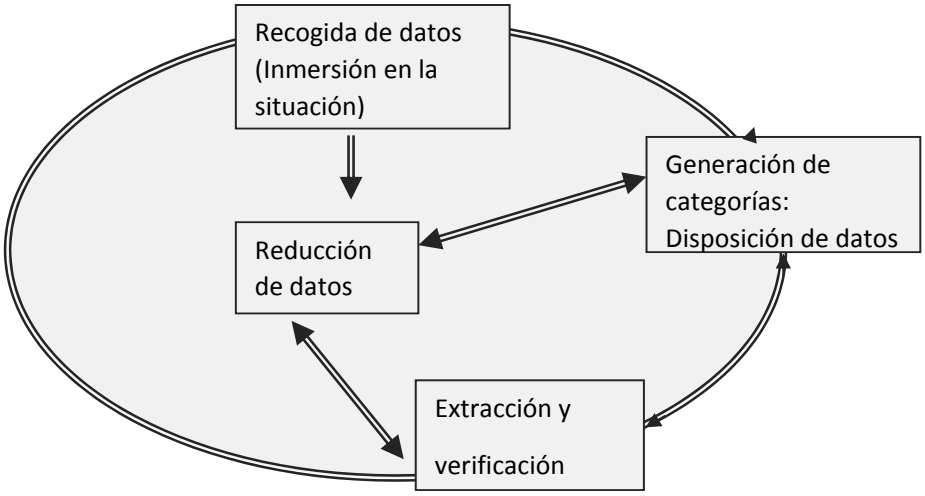
Recogida de datos (Inmersión en la situación)
Generación de categorías: Disposición de datos
Reducción de datos
Extracción y verificación

Figura 1. Componentes del análisis de datos cualitativo.

A las cuatro fases propiamente dichas, se le añade una previa, de delimitación del área de estudio, y otra al final, de presentación de las conclusiones y teorías de apoyo (Figura 2).

Fases del proceso analítico de datos	
1.º	Inmersión en la situación y descubrimiento en progreso de la información y datos. Generación inicial de categorías. Identificación de temas claves y desarrollo de conceptos.
2.º	Con la información recogida, codificación de los datos y el refinamiento de la comprensión del tema de estudio. Selección y definición los conceptos; comparación de los incidentes aplicables a cada categoría.
3.º	Validación de las categorías: Frecuencia de aparición y distribución de los conceptos, se integran las categorías y sus fenómenos.
4.º	Organización, interpretación y explicación de las categorías. Relativización de los descubrimientos, de acuerdo con el contexto en que fueron tomados.

Figura 2. Fases del proceso analítico de datos.

Siguiendo a Taylor y Bogdan (1996), el análisis de los datos cualitativo conlleva los pasos que mencionamos a continuación.

2.1. Descubrimiento en progreso

Los investigadores cualitativos van dando sentido a lo que estudian combinando intuición y perspicacia y familiaridad con los datos. La clave está en buscar temas examinando los datos de todas las maneras posibles, de acuerdo con las siguientes pautas:

1. Lea varias veces su información.
2. Siga la pista de temas, intuiciones, interpretaciones e ideas.
3. Busque los temas claves. Elabore listas posibles de temas claves.
4. Elabore tipologías o esquemas de clasificación útiles para identificar temas y desarrollar conceptos y teorías.
5. Desarrolle conceptos y proposiciones teóricas. Avance de la descripción a la interpretación y a la elaboración de teorías a través de los conceptos y proposiciones claves. En investigación cualitativa los con-

ceptos claves son instrumentos que proporcionan una referencia fundamental, y sugieren direcciones para la observación.

6. Lea bibliografía sobre el tema de estudio. Ayudará a contrastar conceptos, proposiciones y marcos teóricos relevantes a la investigación.
7. Desarrolle una guía del proceso de la investigación. Describa el trabajo en términos generales de modo que oriente el análisis de los lectores.

2.2. Codificación de los datos

En la investigación cualitativa, la codificación es un modo sistemático de desarrollar y refinar las interpretaciones de los datos. El proceso de codificación incluye la reunión y análisis de todos los datos que se refieren a temas, ideas, conceptos, interpretaciones y proposiciones. Taylor y Bogdan sugieren los siguientes pasos para la codificación de los datos cualitativos:

1. Desarrollo de categorías de codificación. Elabore una lista con todos los temas, conceptos, interpretaciones, proposiciones identificados durante el análisis inicial. Asigne un número o letra a cada categoría de codificación.
2. Codifique todos los datos. En las notas de campo indique en el margen el número asignado a cada categoría cuando aparezca en el texto. A medida que se codifican los datos hay que refinar el esquema de la codificación. En algunos casos habrá datos que pertenezcan a más de una categoría, asignándolos a cada una de ellas.
3. Separe los datos pertenecientes a las diversas categorías de codificación conservando intacto el texto original.
4. Vea los datos que han sobrado, que no entran en ninguna de las categorías existentes. Trate de crear nuevas categorías o de situar en las ya existentes. Si no se ajustan no lo fuerce.
5. Refine su análisis volviendo a leer el texto y las categorías. Hay que estar preparado para desarrollar nuevas categorías y prescindir de otras si no las consideramos adecuadas.

2.3. Relativización de los datos y conclusiones

Es la fase final del análisis cualitativo. Se trata de interpretar los datos en el contexto en el que fueron recogidos, teniendo en cuenta las siguientes consideraciones:

1. Datos solicitados o no solicitados. Tratar de ver la información que ha emergido de un modo natural y la que ha forzado el investigador.
2. Influencia del observador sobre el escenario.
3. Influencia de otras personas en el escenario.
4. Datos directos e indirectos. Literales o interpretados.
5. Fuentes. Grupo, informantes clave.
6. Nuestros propios supuestos.

En el trabajo cualitativo el investigador comienza con un mínimo de supuestos aunque nuestros preconceptos son difíciles de evitar. Los datos nunca se explican a sí mismos, sino que son interpretados en el marco de unos supuestos teóricos (Figura 3).

Proceso general de análisis de datos cualitativos		
Tareas	Actividades	Operaciones
Reducción de datos	Separación de unidades	Criterios de separación físicos, temáticos, gramaticales, sociales.
	Identificación y clasificación	Categorización y codificación.
	Síntesis y agrupamiento	Agrupamiento físico, creación de metacategorías, obtención de estadísticos, de agrupamiento y síntesis.
Disposición y transformación de datos	Disposición	Elaboración de tablas, gráficos, modelos, matrices, sistemas de redes.
	Transformación	Expresión de los datos en otros lenguajes (numérico, gráfico).

Proceso general de análisis de datos cualitativos		
Tareas	Actividades	Operaciones
Obtención de resultados y verificación de verificación de conclusiones	Proceso para obtener resultados	Datos textuales: descripción e interpretación; recuento, concurrencia de resultados códigos, comparación y contextualización.
	Proceso para alcanzar conclusiones	Datos textuales: consolidación teórica aplicación de otras teorías, uso de metáforas y analogías, síntesis con resultados de otros investigadores. Datos numéricos: uso de reglas de decisión (comparación de los resultados con modelos teóricos y experiencia del analista).
	Verificación de conclusiones	Comprobación o incremento de mediante la presencia prolongada en la validez el campo, intercambio de opiniones con otros investigadores, triangulación comprobación con los participantes, establecimiento de adecuación referencial, ponderación de la evidencia. Comprobación de la coherencia estructural.

Figura 3. Proceso general de análisis de datos cualitativos (García Jiménez, Gil Flores, Rodríguez González, 1994).

Hay programas informáticos para el análisis de datos cualitativos como el Atlas.ti.

3. CRITERIOS PARA VALORAR LAS INVESTIGACIONES CUALITATIVAS

Los criterios de bondad tienen sus raíces en la concepción del paradigma para el cual se designan, por lo tanto, no se puede esperar aplicar

los criterios positivistas a los estudios naturalistas. ¿Qué criterios podrían servir para el paradigma naturalista? Guba y Lincoln, (1989:236) proponen tres grupos de criterios para valorar las investigaciones naturalistas o constructivistas. El primero de ellos lo consideran paralelo a los criterios clásicos del paradigma positivista de *validez interna, validez externa, fiabilidad y objetividad*. El segundo lo relacionan con los procesos hermenéuticos-dialécticos y el tercero con criterios de autenticidad.

3.1. Criterios de credibilidad, seguridad, confirmabilidad y transferibilidad

Así explican Guba y Lincoln los criterios del paradigma positivista de *validez interna, validez externa, fiabilidad y objetividad* y los paralelos en el paradigma naturalista, interpretativo:

— ***La validez interna*** es la valoración del grado de isomorfismo entre los hallazgos del estudio y el mundo real, y ello no puede estar presente en un paradigma que rechaza el realismo ontológico. Si la realidad sólo existe en las mentes de sus constructores —subjetividad—, ¿qué sentido tiene ese isomorfismo?

— ***La validez externa***, concepto que contiene la esencia de la generalización de os hallazgos de la muestra a la población, igualmente tiene poco significado si las realidades que podríamos desear generalizar existen en forma diferente, en las mentes de cada una de las personas, condicionadas por sus historias y circunstancias diferentes, basadas en sus experiencias particulares e interpretadas desde sus propios sistemas de valores.

— ***La fiabilidad*** es esencialmente una valoración de la estabilidad —del fenómeno que ha sido valorado y del instrumento utilizado para registrarlo—, lo que supone que el fenómeno es invariable, al menos en un corto tiempo. Pero, si el fenómeno cambia, y esto es central en el desarrollo y mejora de las construcciones mentales de las personas, entonces la fiabilidad es inútil como criterio de bondad en este paradigma.

— Finalmente ***la objetividad*** refleja la posición positivista de la dualidad objeto-sujeto, pero si el paradigma rival, el interpretativo, afirma que la interacción (monismo) es inevitable, ¿Qué puede significar la objetividad?

Después de reflexionar sobre los posibles criterios que podrían servir para el paradigma naturalista, proponen un grupo de criterios paralelos a los del paradigma positivista utilizando los términos alternativos de *credibilidad, seguridad —o formalidad, confirmabilidad y transferibilidad—*. Así los explican:

— ***Credibilidad.*** Es un criterio paralelo al de *validez interna*. La atención se centra en que las construcciones de las personas implicadas y las representaciones del propio investigador sean equivalentes. Se utilizan varias técnicas para incrementar la probabilidad de que tal isomorfismo se produzca; son las siguientes:

— Compromiso o encuentro prolongado para evitar efectos de distorsión por la falta de información y falta de confianza para descubrir las construcciones y comprender la cultura del contexto.

— Observación continua para identificar los elementos y las características más relevantes de la situación.

— Informes de colegas a través del compromiso con compañeros desinteresados para la discusión y análisis de las conclusiones de los hallazgos.

— Análisis de los casos negativos y opuestos.

— Seguimiento de la subjetividad con el fin de controlar las construcciones del investigador.

— Comprobaciones de los miembros y participantes de la investigación, que pueden ser formales e informales, cumpliendo un importante número de funciones:

  • Valoración de la intención de las acciones concretas del investigador.

  • Corrección de errores de hechos o de interpretación.

- Posibilidad de información adicional, estimulando la reflexión de los implicados con el fin de confirmar evidencias y acuerdos.
- Síntesis como un primer paso hacia el análisis.
- Valoración de la adecuación de la información posibilitando la confirmación de los datos individuales por parte de los implicados.

— ***Seguridad o formalidad.*** Se corresponde con el concepto de *fiabilidad* y está relacionado con la estabilidad de los datos en el tiempo. En el paradigma naturalista los cambios producidos por modificaciones metodológicas pueden contribuir a mejorar la seguridad de la información en un proceso de triangulación.

— ***Confirmabilidad.*** Es un criterio paralelo al de *objetividad*, interesado en asegurar que los datos, interpretaciones y logros de la investigación tengan sus raíces en los contextos y en las personas. Esto significa que los datos, y la lógica usada en su interpretación, deben estar unidos a sus fuentes corroborando la totalidad explícita e implícita en un informe narrativo de estudio de caso. La técnica utilizada frecuentemente para confirmar los datos y su interpretación es la auditoría de confirmabilidad, cuyas raíces se encuentran en la auditoría fiscal.

— ***Transferibilidad.*** Corresponde al criterio de *validez externa o generalización*. Consiste en un proceso empírico para evaluar el grado de similitud entre los contextos. A diferencia de la generalización en el paradigma positivista, aquí se trata de una cuestión relativa y la técnica adecuada consiste en hacer una descripción detallada del tiempo, lugar, contexto y cultura de la investigación.

3.2. Criterios hermenéuticos

Los criterios hermenéuticos son utilizados en el paradigma naturalista para enjuiciar sus investigaciones. El proceso hermenéutico-dialéctico se utiliza a modo de un control de calidad. Los datos son analizados inmediatamente a su producción con la finalidad de realizar una

comprobación mediante comentarios, revisiones, ampliaciones, o cualquier otro modo que los participantes proporcionen durante las reuniones. Se buscan los diferentes niveles de significado en el significado aparente, los diferentes significados subyacentes al significado literal y son incorporados a medida que emergen, en un proceso de reconstrucción colaborativa. Cómo método de análisis la hermenéutica sugiere una manera de entender los datos textuales, dado que se centra en el significado del texto.

3.3. Criterios de autenticidad

El tercer grupo de criterios que utiliza la investigación constructivista para enjuiciar su calidad es el de la *autenticidad*. Los dos criterios anteriores no son de entera satisfacción para los constructivistas; los primeros son, principalmente, criterios metodológicos; las metas, los resultados y la negociación son criterios igualmente importantes para juzgar una investigación. De modo que, basarse sólo en criterios metodológicos sería insuficiente para garantizar el esfuerzo investigador para lograr algo —resolver un problema—.

El segundo grupo de criterios —el proceso hermenéutico tiene más que ver con los procesos constructivistas pero es débil por estar implícito en el proceso mismo. Por ello, han ideado los criterios de autenticidad que emergen de la concepción básica del constructivismo. Los criterios de autenticidad, representan el tercer grupo de criterios para valorar las investigaciones, e incluyen los siguientes elementos (Lincoln y Guba, 1986:81):

— ***Justicia.*** Se refiere a la estructura de valores subyacentes en el proceso de investigación. El papel del evaluador es el de indagar para comunicar dichos valores y ver los que están en conflicto a través de la diferencia entre las creencias y el sistema de valores, mediante la realización de una negociación de recomendaciones abierta y una agenda de las acciones subsiguientes.

— ***Autenticidad ontológica.*** Tiene que ver con como maduran, mejoran, se expanden y se elaboran las construcciones de los partici-

pantes, dado que han alcanzado mayor información y han logrado un uso más sofisticado de ella. Es, literalmente, «una mejora en el individuo —o grupo— de la conciencia experiencial del mundo». La autenticidad ontológica se refuerza mediante la experiencia vicaria y se comprueba por el testimonio de informantes seleccionados, además de por una auditoría judicial para el estudio del caso determinado.

— ***Autenticidad educativa.*** Representa el grado en el que los participantes individuales comprenden y aprecian las construcciones de los otros fuera de su grupo particular.

— ***Autenticidad catalítica.*** Consiste en el grado en el cual la acción es estimulada y facilitada mediante el proceso de la investigación. El propósito de la investigación es la acción y la toma de decisiones, y no sería completa sin una implicación en la acción por parte de los participantes.

— ***Autenticidad táctica.*** Se refiere al poder de los participantes para actuar.

La autenticidad debería ser el criterio más importante para valorar la investigación naturalista. Y es que, como ya hemos visto, la autenticidad integra conceptos como el de justicia, conciencia, enriquecimiento, apreciación de los puntos de vista y necesidades de los otros, estimulación hacia la acción, y generación de cambios deseables.

4. EL ANÁLISIS DE DATOS EN LA INVESTIGACIÓN CUANTITATIVA

4.1. Distribuciones numéricas y reducción de datos

Las puntuaciones obtenidas mediante los instrumentos de recogida de información en una determinada variable se denominan 'puntuaciones directas', y se representan mediante el símbolo Xi. Una primera aproximación a su estudio nos la aporta su presentación en tablas de datos, pudiendo mostrarse agrupadas por frecuencias, porcentajes o por intervalos.

Tabla 1. Tabla de datos por frecuencias y porcentajes

Xi	fi	%
3	2	11,70
4	2	11,70
5	3	17,60
6	4	23,00
7	3	17,00
8	2	11,70
9	1	5,80
	n = 17	100,00

4.2. Niveles de medida y tipos de variables

Dependiendo de las posibilidades de asignar códigos numéricos a las características que se observan en las personas u objetos, las variables se clasifican en nominales, ordinales, de intervalo y de razón.

El nivel nominal es el más simple en cuanto a la asignación de números a las variables observadas, permitiendo clasificar los individuos y objetos en clases o categorías descriptivas que deben ser exhaustivas en sí mismas y excluyentes. El nivel ordinal además de clasificar ordena en función de que posea en mayor o menor grado la característica que se pretende medir. El nivel de intervalo, clasifica, ordena, y además especifica las distancias existentes entre las distintas categorías. El nivel de razón o de proporción, incluye además, el 0 absoluto, lo que representa ausencia total de la característica que se mide.

Las variables medidas al nivel de intervalo y de razón son variables continuas, como por ejemplo, las puntuaciones obtenidas por un grupo de alumnos en un test de inteligencia, el tiempo tardado por otro grupo en la realización de una tarea, las calificaciones obtenidas en una prueba precisa de rendimiento, reflejando, todas ellas, un conjunto ordenado de valores, pudiendo adoptar infinitos valores fraccionados, o intermedios.

Esto tipo de datos adoptan una distribución continua y permiten operar con ellas con pruebas estadísticas propias de los datos medidos a nivel de intervalo y de razón.

Por otro lado, las variables medidas al nivel ordinal son variables discretas, las distancias entre ellas no son equidistantes, adoptando una ordenación de grado, o de orden, de menor a mayor, o al revés. Pueden mostrar una continuidad latente en la variable, y en estos casos se pueden analizar considerándolas variables continuas. Estas variables suelen ser categorías, por ejemplo, si nos referimos a un grupo de personas estudiadas atendiendo a su nacionalidad, las clasificamos en función de categorías de nacionalidades, como por ejemplo: española, francesa, inglesa, o si las estudiamos en función de su nivel de estudios, tendremos las categorías de estudios primarios, secundarios, superior.

4.3. La distribución continua y la distribución normal de probabilidad

Cuando las variables aleatorias que medimos son continuas, o básicamente continuas, su representación gráfica puede hacerse mediante histogramas, y cuando hay una gran densidad de valores de la variable medida tienen una forma de curva bien delineada. Cuando la distribución es muy grande la curva adopta la forma de curva normal o distribución normal. La distribución normal es la distribución de probabilidad más útil e importante en estadística, es simétrica y tiene forma de campana.

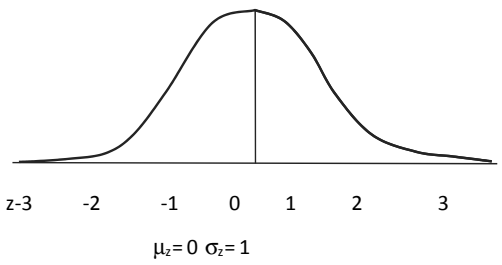

Figura 4. Distribución de la curva normal.

La distribución normal se denomina z, y constituye la base de un gran grupo de pruebas estadísticas para la inferencia. La probabilidad equivalente al área comprendida entre la curva normal y el eje de abscisas es igual a 1.

Cualquier variable aleatoria puede transformarse a la forma estándar de la distribución de la curva normal. A esta transformación de la variable aleatoria se le denomina tipificación de la variable, y la nueva transformación de la variable sufre una transformación lineal de la inicial mediante la fórmula de z, equivalente a la puntuación directa menos la media dividida por la desviación típica:

$$z = \frac{X_i - \mu}{\sigma}$$

4.4. Representaciones gráficas

Las representaciones gráficas aportan información valiosa que en muchas ocasiones no afloran en los índices estadísticos. El tipo de gráfica que elijamos para la representación de los datos dependerá de las características de la distribución de la variable, de si es continua o discreta, si ha venido medida por instrumentos que alcanzan el nivel de medida de intervalo, ordinal o nominal, y si las puntuaciones son directas o acumuladas. Así, el diagrama de barras (Figura 5), representa cada valor mediante una barra proporcional a la frecuencia obtenida por dicho valor en la variable, siendo ésta una variable discreta, no continua, por lo que las barras están separadas. La base de la barra adopta un ancho igual para todas las barras, elegido por el investigador, y la altura viene determinada por la frecuencia obtenida en ese valor.

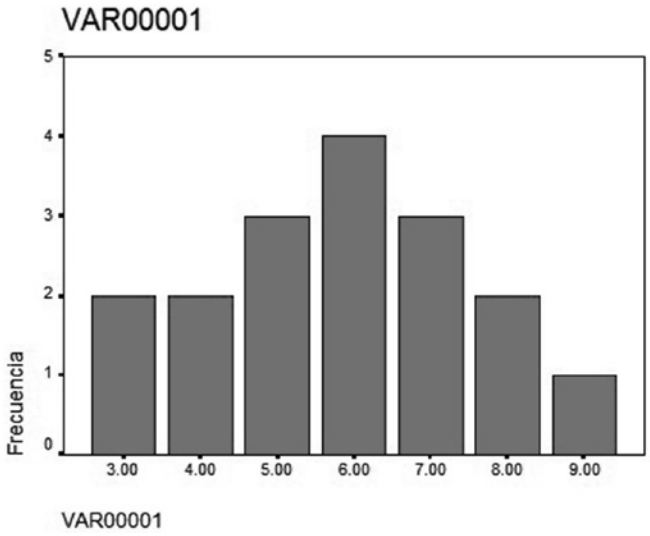

Figura 5. Diagrama de barras.

Para la representación de variables continuas, agrupadas en intervalos, se utiliza el histograma (Figura 6). En él, por tratarse de una variable continua, las barras van unidas, pudiendo mostrar, también la curva normal. La altura de la barra indica la frecuencia correspondiente al intervalo o valor de la variable X, y en la base se indica el valor medio del intervalo.

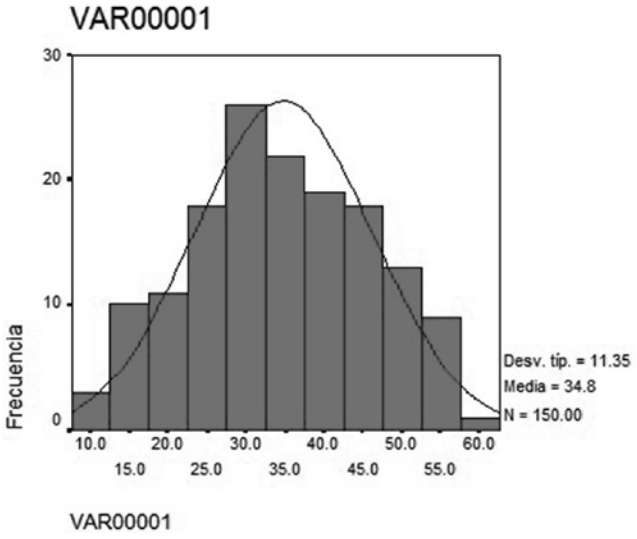

Figura 6. Histograma con curva normal.

El diagrama circular de sectores (Figura 7) se utiliza para representar variables discretas, aunque también puede utilizarse con variables continuas. La amplitud de los sectores es proporcional a su frecuencia.

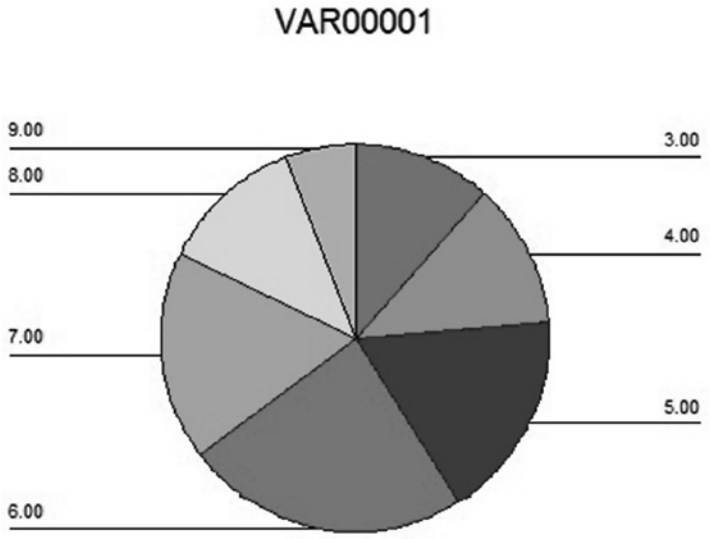

Figura 7. Diagrama de sectores.

4.5. Medidas de tendencia central. La moda, la mediana y la media

Las medidas de tendencia central informan de los valores más representativos del conjunto de una serie de datos, y suelen venir representados por la moda, la mediana y la media de la distribución de la serie de datos.

La moda (Mo) de una serie de datos se define como el valor más frecuente, el que predomina.

La mediana, en una serie de datos ordenados de menor a mayor, es el valor del dato que ocupa el lugar intermedio, y deja por debajo y por encima de sí el 50% de los casos.

La media aritmética es el promedio de todos los valores obtenidos por un grupo en una variable, y se calcula sumando todas las puntuaciones de la distribución dividiéndola por el total de puntuaciones. Es la

medida de tendencia central más usual, y su utilización requiere operar con datos medidos en nivel de intervalo o de razón. Su fórmula es:

$$\overline{X} = \frac{\sum X_i}{n}$$

Siendo:

SX_i : la suma de las puntuaciones directas.
n: el número de casos (sujetos si se trata de un grupo de personas).

La mediana coincide con la media y con la moda en una distribución normal.

4.6. Medidas de posición. Los cuartiles, los deciles, los percentiles

Hemos dicho que la mediana representa el valor central de la serie de datos, dejando por debajo y por encima de dicho valor el 50% de los casos de la serie de datos u observaciones de la distribución. A partir de la mediana, se pueden establecer distintas posiciones en la serie, con el fin de conocer la posición que ocupa una determinada puntuación dentro de la serie de datos.

Si la serie la dividimos entre cuatro a cada una de las partes la denominamos cuartil, si es entre diez, a cada parte la denominamos decil y si es entre cien, percentil.

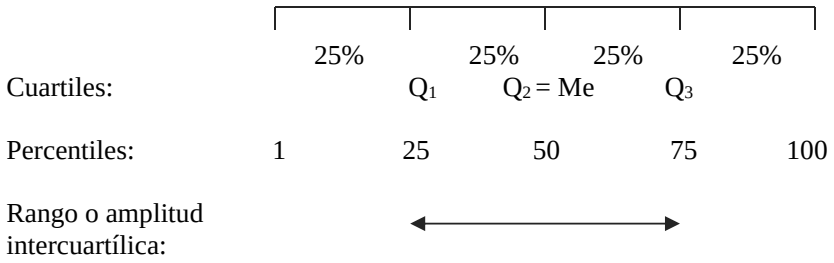

Figura 8. Cuartiles, percentiles y amplitud intercuartílica.

La amplitud intercuartílica es la diferencia entre el cuartil inferior y el cuartil superior, tomando en consideración sólo el 50% de las puntuaciones centrales de la serie.

4.7. Medidas de variabilidad. La amplitud, la desviación media, la varianza y la desviación típica

Las medidas o índices de variabilidad o de dispersión dan idea del alejamiento de los datos grupales respecto de las medidas de centralización. Los índices de variabilidad informan sobre el grado de concentración o dispersión de los valores de la distribución en torno al valor central. Si el índice está cercano al promedio se dice que la distribución es homogénea; si está lejano se dice que es heterogénea.

Entre las medidas de variabilidad o dispersión tenemos la amplitud, la desviación media, la varianza y la desviación típica.

La amplitud total (A_T) de un conjunto de datos, es la diferencia entre el valor más pequeño y el más grande de la serie de datos, y representa el recorrido de la variable. Se halla restando a la puntuación final (X_f), o mayor, la puntuación inicial (X_i), o menor, sumándole una unidad.

$$A_T = X_f - X_i + 1$$

La desviación media se basa en la desviación de cada uno de los valores que toma la variable con relación a la media.

$$DM = \frac{\left|X_1 - \bar{X}\right| + \left|X_2 - \bar{X}\right| + ... + \left|X_n - \bar{X}\right|}{n} = \frac{\sum \left|X_i - \bar{X}\right|}{n}$$

La desviación media (DM) con respecto a la media es igual a la suma de cada una de las puntuaciones directas menos la media, dividida entre el número total de sujetos o puntuaciones de la serie, en valores absolutos.

La varianza (s^2) pretende resaltar la desviación de las desviaciones importantes, las más alejadas de la media, sobre las pequeñas, mediante la elevación al cuadrado de todas las desviaciones. Para su cálculo, con puntuaciones directas, se suele utilizar la siguiente fórmula:

$$s^2 = \frac{\sum X_i^2}{n} - \bar{X}^2$$

Al tratarse de un valor obtenido en función de las distancias cuadráticas de cada puntuación de la distribución con respecto a la media aritmética, adopta siempre valores positivos. La varianza es una medida estadística muy utilizada para conocer las diferencias existentes entre dos o más muestras.

La desviación típica (s) es la raíz cuadra de la varianza:

$$s = \sqrt{s^2}$$

Con muestras de tamaño pequeño, n< 30, al denominador se le resta una unidad. Con puntuaciones agrupadas y n<30, la fórmula es la siguiente:

$$s = \sqrt{\frac{\sum x_i^2 n_i}{n-1}}$$

Siendo:

$$x_i = X_i - \overline{X}$$

La desviación típica es la medida de variabilidad más utilizada, por ser la más fiable, en el sentido de que es la que menos varía de una muestra a otra. La principal ventaja de la desviación típica sobre la varianza es que la desviación típica viene dada en las mismas unidades que la media. Como medida descriptiva se utiliza más la desviación típica que la varianza.

4.8. Puntuaciones individuales. Las puntuaciones diferenciales y puntuaciones las típicas

A partir de las puntuaciones directas en una serie de datos de un grupo, se puede estudiar la relación entre la puntuación directa de un sujeto, X_i, con las del grupo. Así, tenemos la puntuación diferencial, que es la diferencia entre una puntuación directa individual X_i y la media aritmé-

tica del grupo, e indica la desviación de dicha puntuación con respecto a la media grupal. Su fórmula es la siguiente:

$$x_i = X_i - \bar{X}$$

Las puntuaciones diferenciales de dos series de datos distintas podrían coincidir en su valor pero no significar lo mismo, no permitiendo comparaciones entre ellas. Las puntuaciones típicas aportan una solución a esta dificultad, definiéndose la puntuación típica como la diferencia entre una puntuación directa en una variable y su media (puntuación diferencial) dividida entre la desviación típica de esa misma variable.

$$z_i = \frac{x_i}{s_x} = \frac{X_i - \bar{X}}{s_x}$$

La puntuación típica indica el número de desviaciones típicas que una determinada puntuación directa se aparta de la media del grupo. La media de la distribución de las puntuaciones típicas z es igual a cero. La varianza y la desviación típica de la distribución de las puntuaciones típicas z es igual a uno, 1.

Las puntuaciones típicas reflejan las relaciones entre las puntuaciones directas y las grupales de una serie de datos, independientemente de la unidad de medida utilizada, lo que permite la comparación entre series de datos distintas.

5. LAS MEDIDAS CORRELACIONALES BIVARIADAS Y SU INTERPRETACIÓN

Las distribuciones bidimensionales son el resultado de las observaciones realizadas de manera simultánea sobre dos características de una determinada muestra o de una población. A cada uno de los casos, o sujetos observados, le corresponden dos observaciones que pueden ser dos valores numéricos, si se trata de características cuantificables, o dos categorías, si se trata de variables categóricas o discretas.

Una relación queda establecida por un conjunto de pares ordenados, en un orden fijo. Los valores observados en las variables se representan

mediante una tabla de doble entrada. Gráficamente pueden representarse mediante los ejes cartesianos. Los puntos de cruce entre ambos valores son los pares ordenados de los valores de ambas variables.

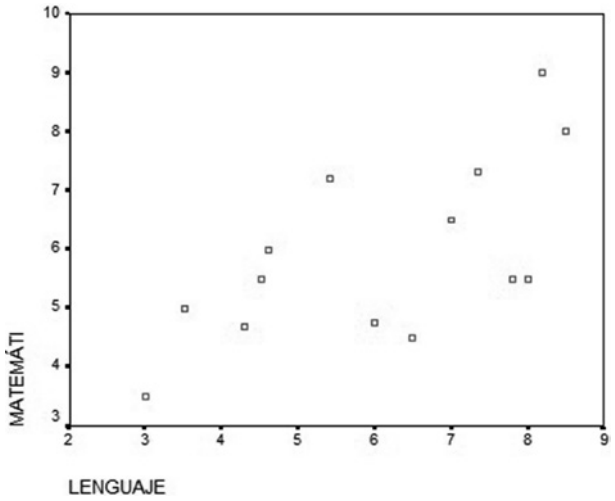

Figura 9. Diagrama de dispersión: Relaciones entre X e Y.

La correlación es una medida del grado de relación entre dos variables. Los coeficientes de correlación aportan información sobre la intensidad y la dirección con que dos variables están relacionadas y permiten establecer comparaciones entre diversos casos. Estos coeficientes se basan en la variación concomitante de los miembros del conjunto de pares ordenados. Si covarían, es decir, varían juntos y en forma correspondiente, por ejemplo valores altos con valores altos, valores bajos con valores bajos, se dice que hay correlación positiva. Si covarían valores altos con valores bajos, entonces se dice que hay correlación negativa. La correlación bivariada, por tanto, expresa el grado y sentido de la variación concomitante de dos series de datos.

Dependiendo de que las observaciones hayan sido medidas en escalas de intervalos, ordinales o categoriales, el coeficiente utilizado para calcular el grado de relación existente entre ambas variables será distinto. Así, utilizaremos el coeficiente de correlación de Pearson para las variables que alcanzan nivel de medida de intervalo, el coeficiente de Spearman para el nivel de medida ordinal, y el de «chi cuadrado» y contingencia para el nivel de medida categorial.

5.1. La correlación producto momento de Pearson

La correlación de Pearson, o correlación producto-momento, (r_{xy}, correlación de la variable X con la variable Y), entre dos variables medidas en escala de intervalo, se define como el cociente entre la covarianza entre las variables X e Y, dividida entre el producto de las desviaciones típicas de cada una de las variables:

$$r_{xy} = \frac{s_{xy}}{s_x s_x}$$

La ecuación de la correlación de Pearson, r_{xy}, para puntuaciones directas, es:

$$r_{xy} = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{\left[n\sum X^2 - (\sum X)^2\right] \times \left[n\sum Y^2 - (\sum Y)^2\right]}};$$

El índice correlacional nos permite realizar comparaciones entre el grado de relación existente entre distintos casos y grupos de sujetos, y conocer la intensidad de la relación. Así, por ejemplo, en las dos series de datos siguientes:

Tabla 3. Tabla de correlación

Serie A	Serie B
0	12
1	19
2	23
3	34
4	56
5	62

La serie A representa la variable X_i y la serie B representa la variable Y_i. El cálculo de la correlación nos da un valor r_{xy} = 0,97, muy próximo

a 1, lo que nos indica que existe una correlación muy elevada entre ambas series, casi perfecta. Su interpretación mediante el análisis de datos estadísticos informáticos con programas cono el SPSS, nos vendría indicada mediante los niveles de significación alfa (a).

Tabla 4. Estadísticos descriptivos

	N	**Media**	**Desviación típica**
Variable X	6	2.500	1.8708
Variable Y	6	34.333	20.4808

Tabla 5. Correlaciones, suma de cuadrados y productos cruzados y covarianzas

		Variable X	**Variable Y**
Variable X	Correlación de Pearson	1	.971(**)
	Sig. (bilateral)	.	.001
	Suma de cuadrados y productos cruzados	17.500	186.000
	Covarianza	3.500	37.200
	N	6	6
Variable Y	Correlación de Pearson	.971(**)	1
	Sig. (bilateral)	.001	.
	Suma de cuadrados y productos cruzados	186.000	2097.333
	Covarianza	37.200	419.467
	N	6	6

** La correlación es significativa al nivel 0,01 (bilateral).

5.2. La correlación ordinal de rangos de Spearman

Cuando se trabaja con variables medidas con un nivel ordinal, significa que las puntuaciones obtenidas mediante un cuestionario, o una escala de observación, conforman una sucesión de valores ordenados de mayor a menor. En estos casos el coeficiente de correlación a aplicar es el de *Spearman* cuya fórmula es la siguiente:

$$r_s = 1 - \frac{6\sum D^2}{n(n^2 - 1)};$$

siendo D: la diferencia de rango entre la variable X y la variable Y, y n: número de casos

Veamos un ejemplo. Queremos conocer si existe alguna relación entre las variables «claridad expositiva» de un profesor de Matemáticas, que la nombramos como X_i, y su «capacidad para motivar a los alumnos», que la nombramos como Y_i. Para medir cada una de las variables, se ha utilizado un cuestionario de respuesta escalada, entre 1, mínima y 5, máxima, que ha sido contestado por 12 alumnos. Las puntuaciones obtenidas son las siguientes:

Tabla 6. Tabla de correlación

Alumno	Variable Xi	Variable Yi
1	50	43
2	47	50
3	45	46
4	40	40
5	40	40
6	35	32
7	32	37
8	27	35
9	25	28
10	25	30
11	22	22
12	19	26

El valor del coeficiente de correlación a partir de los datos, $r_s = 0,95$, es elevado, cercano a 1, lo que indica la existencia de relación muy alta entre ambas variables. Mediante el programa de análisis estadístico de datos SPSS, obtenemos la siguiente tabla de resultados:

Tabla 7. Tabla de resultados de la correlación Rho de Spearman.

		Variable X	**Variable Y**
Variable X	Coeficiente de correlación	1.000	.949(**)
	Sig. (bilateral)	.	.000
	N	12	12
Variable Y	Coeficiente de correlación	.949(**)	1.000
	Sig. (bilateral)	.000	.
	N	12	12

** La correlación es significativa al nivel 0,01 (bilateral).

6. COMPARACIÓN ENTRE GRUPOS. DISEÑOS DE GRUPO

6.1. Concepto de experimento y de diseño de grupo experimental

Un experimento consiste en modificar deliberadamente y de modo controlado las condiciones que determinan un hecho y en observar e interpretar los cambios que ocurren en él. El objetivo de los diseños experimentales es comprobar los efectos de un tratamiento, llamado variable experimental, o independiente— sobre la variable dependiente o resultado.

A. Los diseños de grupo experimental ofrecen una estructura seleccionada por el investigador para responder a las preguntas de investigación en condiciones tales que garanticen la credibilidad y validez de los resultados obtenidos.

B. Para verificar que un determinado tratamiento, o programa, mejora un resultado, requiere hacer comparaciones entre las medidas de los individuos de varios grupos, de los cuales, al menos uno de los grupos recibe el tratamiento o el nuevo programa, y al menos en otro de los grupos no se introduce ninguna modificación en el tratamiento o programa. El primer grupo recibe el nombre de grupo experimental y el otro recibe el nombre de grupo de control o de referencia. El resto de condiciones que puedan influir en los resultados deben permanecer tan homogéneas como sea posible en todos los grupos, para garantizar, en términos de probabilidad, que existen relaciones de causalidad entre las diferencias de resultados obtenidas en los grupos debidas al tratamiento. (Campbell & Stanley, 1963; Cook & Campbell, 1979; Kerlinger, 1973).

C. El modo de comparar las puntuaciones obtenidas por los grupos, para probar si existen diferencias debidas al tratamiento, se puede hacer por medio de la prueba 't de Student' o la prueba 'F' del análisis de la Varianza, entre otras muchas posibles.

D. El procedimiento clásico de prueba de hipótesis sigue los siguientes pasos:

1. Formulación de las hipótesis, nula y alternativa (H_0, H_1).
2. Seleccionar el nivel de significación (alpha).
3. Determinar la prueba de hipótesis a utilizar (z, t, χ^2, F, T, U).
4. Definir la región de rechazo (pruebas de hipótesis de dos colas, o de una cola).
5. Definir las normas de decisión para rechazar la hipótesis nula.
6. Realizar los cálculos necesarios sobre los datos obtenidos tras el periodo de aplicación del tratamiento en la variable resultados o dependiente, tanto en el grupo experimental, en el que se ha aplicado el nuevo tratamiento, o programa, como en el grupo de control, en el que no se han introducido cambios en el programa, y realizar la prueba estadística para probar las hipótesis.
7. Comparar el valor obtenido en la prueba de hipótesis con la norma de decisión, y tomar la decisión acerca del posible rechazo de la hipótesis

nula. Si se rechaza la hipótesis nula (que dice no existen diferencias estadísticamente significativas entre los resultados obtenidos en los grupos) se puede confirmar la hipótesis alternativa.

6.2. Resolución de un diseño investigación de dos grupos independientes y nivel de medida de intervalo

Veamos el siguiente ejemplo: Un investigador evaluativo quiere probar la eficacia pedagógica de los cursos formativos a través de internet, mediante programas que utilizan diseños instructivos con ayudas al aprendizaje del alumno como orientaciones para el estudio, autoevaluaciones, guía para la realización de problemas y actividades prácticas, interactividad entre profesor y alumno, y entre los alumnos, frente a otros programas que no introducen estas ayudas. Selecciona dos grupos de alumnos, de forma independiente, pertenecientes a la misma población, concretamente del segundo curso de Educación Social de la UNED, con las asignaturas de Metodología de Investigación en Educación aprobadas, lo cual supone que tienen la misma varianza, son homogéneos, o no difieren significativamente en la variable que vamos a estudiar, el rendimiento en un curso sobre Metodología de investigación. La muestra invitada a participar es de 60 alumnos y, a partir de ella, formamos dos grupos asignados al azar a cada uno de los grupos.

El grupo experimental, formado por 30 alumnos, utilizará un programa de estudio de los temas de inferencia estadística, a través de internet, con ayudas al aprendizaje. El grupo de control, formado por los otros 30 alumnos seleccionados, utilizará el mismo programa de estudio, a través de internet, pero sin elementos de ayuda al aprendizaje. Se dispone de mes y medio para comprobar los resultados de ambos enfoques, y al finalizar dicho periodo, de acuerdo con la información y normas conocidas por los estudiantes, se envía a los alumnos, en un día y hora previamente fijados, a través de la plataforma educativa, una prueba precisa para comprobar el aprendizaje de los contenidos del programa, y que debe ser resuelta y subida a la plataforma en un máximo de tiempo de hora y media. Por diversos motivos los alumnos que contestan a tiempo son 12 del grupo experimental y 10 del grupo de control. Una vez corregidas las pruebas nos aportan los siguientes datos:

Tabla 8. Resultados en la prueba de aprendizaje

Grupo Experimental	Grupo Control
5	4
5	4
5	5
6	5
6	6
7	6
7	6
7	7
8	7
8	7
9	
9	

Para probar la incidencia del método apoyándonos en los resultados muestrales se procede del siguiente modo:

1.º Esquema y comentario del diseño planteado.

Tabla 9. Esquema del diseño de investigación

Grupos	Selección	Asignación	Pretest	Variable Independiente	Postest
Control	Azar	Aleatoria	No	Programas sin ayudas	Sí
Experimental	Azar	Aleatoria	No	Programas con ayudas	Sí

Se trata de un diseño de investigación cuasi-experimental, cuyos sujetos han sido seleccionados al azar, y el tratamiento se ha asignado de modo aleatorio a los grupos, siendo el que lo recibe el grupo experimental y el otro hace de control del primero. Las varianzas de la población son desconocidas pero supuestas iguales para ambos grupos, que proceden de la misma población. El tamaño de los grupos es pequeño, y con n diferente:

$n_A = 12$ y $n_B = 10$. Sin aplicarles pretest previo. La variable independiente o tratamiento se aplica en dos niveles, el programa estudiado a

través de internet sin ayudas y con ayudas didácticas. La variable dependiente, el rendimiento, se mide mediante una prueba de aprendizaje precisa que alcanza el nivel de medida de intervalo.

2.º Formulamos las hipótesis, tanto literales como las hipótesis estadísticas. Hipótesis textuales:

H_0: El aprendizaje a través de internet, mediante programas educativos específicos, no incidirá en el rendimiento de los alumnos.

H_1: El aprendizaje a través de internet, mediante programas educativos específicos, incidirá en el rendimiento de los alumnos.

Hipótesis estadísticas:

$$H_0 = \overline{X} = \mu$$
$$H_1 = \overline{X} \neq \mu$$

El nivel de significación fijado para el contraste de hipótesis es a = 0,05.

Se trata de una hipótesis bilateral, debido a que el investigador no se decanta abiertamente por una incidencia positiva que le llevaría a formular una hipótesis unilateral del tipo:

$$H_1 = \overline{X} \geq \mu$$

3.º La prueba utilizada para la recogida de los datos. En nuestro caso, para valorar el rendimiento final de los alumnos de ambos grupos, hemos utilizado una prueba precisa de acierto o error. La puntuación total es la suma de las preguntas contestadas acertadamente. El nivel de medida se considera, por tanto, de intervalo, o próximo.

4.º La elección de la prueba estadística para el contraste de hipótesis. De acuerdo con la información comentada, la prueba estadísticas es la t de Student, cuya fórmula es la siguiente:

$$t = \frac{\left|\overline{X}_A - \overline{X}_B\right|}{\sqrt{\left[\frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2}\right] \times \left[\frac{1}{n_A} + \frac{1}{n_B}\right]}}$$

5.º Aplicación de la prueba estadística para el contraste de hipótesis. Realizamos los análisis mediante el programa informático SPSS, aportándonos las siguientes tablas de resultados:

Tabla 10. Estadísticos descriptivos de grupo

GRUPO	N	Media	Desviación típ.	Error típ. de la media
1= Experimental	12	6,3333	1,614	0,466
2= Control	10	5,6000	1,349	0,426

En los estadísticos descriptivos se observa que la media del grupo experimental es superior a la del grupo de control.

Tabla 11. Prueba t para muestras independientes

Prueba de muestras independientes		Prueba de Levene para la igualdad de varianzas		Prueba t para la igualdad de medias						
		F	Sig.	t	gl	Sig. Bilateral	Difere. de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia	
									Inferior	Superior
PUNTUACIONES	Asumen varianzas iguales	0,58	0,455	1,14	20	0,267	0,7333	0,6427	-0,607	2,076
	Asumen varianzas diferentes			1,16	19,9	0,260	0,7333	0,6319	-0,585	2,052

La prueba de Levene, de homogeneidad de las varianzas, muestra que no existen diferencias significativas en las varianzas ($p = 0,455 > 0,05$), lo que se interpreta como que ambos grupos proceden de la misma población, en relación con la variable que se estudia. Y el valor crítico o de significación para la hipótesis bilateral es muy alto, $p = 0,267 > 0,05$, lo que corresponde a un nivel de confianza del 76,3%, no aceptable científicamente para rechazar la hipótesis de nulidad. Los resultados obtenidos tras el contraste de hipótesis. Indican que el rendimiento obtenido por el grupo que recibe ayudas durante el proceso de aprendizaje no ha sido estadísticamente significativamente mayor que el que no las recibe.

Estos resultados aconsejan repetir la experiencia, con una mejora de las ayudas al aprendizaje, junto con un aumento importante del tamaño de la muestra, para asegurar que las muestras productoras de información y datos sean superiores a 30 alumnos en cada uno de los grupos, lo que nos permitirá hacer mejorar la validez interna y externa del estudio.

6.3. Resolución de un diseño investigación de dos grupos relacionados y nivel de medida de intervalo

El equipo docente de Ciencias está preocupado por el bajo rendimiento de los alumnos y quiere probar la eficacia de un nuevo método de enseñanza para comprobar su incidencia en el rendimiento de los alumnos. Cuenta con dos grupos de clase de 25 y 27 alumnos cada uno. Asigna al azar el método que quiere probar a uno de los grupos, continuando en el otro con el que venía utilizando hasta ahora. Pasado un trimestre de la experiencia aplica a todos los alumnos una prueba objetiva de rendimiento bastante precisa, y para realizar el análisis de los resultados va a trabajar sólo con los resultados de las diez parejas de alumnos que obtuvieron en una prueba al inicio del curso los resultados más igualados; cada miembro de la pareja procede de uno de los grupos.

Como vemos, el tratamiento ha sido asignado al azar, pero los alumnos para el análisis de los datos han sido seleccionados de acuerdo con el rendimiento obtenido en una prueba previa en la asignatura de Ciencias Naturales, tomando las 10 parejas cuyas puntuaciones están mas igualadas, ejerciendo de este modo un control sobre posibles variables extrañas

que puedan interferir en nuestro experimento, como son los conocimientos previos en la materia sobre la que estamos investigando.

Pasamos a resolver el diseño de investigación.

1.º Esquema y comentario del diseño planteado.

Tabla 12. Esquema del diseño de investigación

Grupos	Selección	Asignación	Pretest	Variable Independiente	Postest
A= Control	Emparejamiento	Azar	No	Método tradicional	T1
B= Experimental	Emparejamiento	Azar	No	Método nuevo	T1

Se trata de un diseño cuasi-experimental de dos grupos relacionados, cuya selección de sujetos se ha hecho mediante una prueba previa para emparejarlos, la asignación de los tratamientos se ha realizado al azar, sin pretest, con una sola variable independiente, el método de aprendizaje, con dos niveles, el que se venía utilizando y el nuevo.

2.º Formulación de las hipótesis. Hipótesis de nulidad textual:

H_0 : El método de aprendizaje no influye en los resultados académicos.

Estadísticamente: $\mu_A = \mu_B$

Hipótesis alternativa textual:

H_1 : El nuevo método de aprendizaje producirá mejores resultados académicos que si se sigue con el método tradicional.

Estadísticamente: $\mu_A \neq \mu_B$

El nivel de significación elegido es a= 0,01.

Se trata de un contraste unilateral izquierdo, dado que se espera conseguir mejorar el rendimiento con el nuevo método de enseñanza.

3.º Justificación de la prueba elegida para el contraste de hipótesis. Por tratarse de un diseño de dos grupos relacionados, y los datos de

la variable dependiente medidos por el postest alcanzan un nivel de medida de intervalo (rendimiento medido a través de una prueba objetiva precisa), y supuestas las condiciones de normalidad y homogeneidad de las varianzas, elegiremos una prueba paramétrica, en este caso para muestras pequeñas y grupos relacionados, la 't de Student', cuya fórmula es la siguiente:

$$t = \frac{\left|\overline{X}_A - \overline{X}_B\right|}{\sqrt{\frac{\sum D^2 - \frac{(\sum D)^2}{n}}{n(n-1)}}}$$

4. Aplicación de la prueba estadística a los datos. Los resultados obtenidos por las parejas de alumnos en la prueba de rendimiento con la que se ha medido la variable dependiente, son los que se presentan en el siguiente cuadro de datos:

Tabla 13. Resultados en la prueba de rendimiento

Parejas	Grupo de Control	Grupo Experimental	D	D2
1	8	7	1	1
2	6	8	-2	4
3	5	6	-1	1
4	2	4	-2	4
5	3	4	-1	1
6	6	5	1	1
7	7	9	-2	4
8	6	5	1	1
9	4	6	-2	4
10	8	9	-1	1
Totales	55	63	-11+3= -8	22

En la fórmula de t para grupos correlacionados sustituimos los valores de la tabla:

$$t = \frac{\left|\overline{X}_A - \overline{X}_B\right|}{\sqrt{\frac{\sum D^2 - \frac{(\sum D)^2}{n}}{n(n-1)}}} = \frac{|5,5-6,3|}{\sqrt{\frac{22-\frac{64}{10}}{10(10-1)}}} = \frac{0,8}{\sqrt{\frac{15,6}{90}}} = \frac{0,8}{0,416} = 1,92$$

5.º Interpretación del valor empírico obtenido en la prueba estadística y conclusiones. Contrastamos el valor de t empírico, 1,92, con el valor de t crítico de las tablas de la distribución de t, para hipótesis de una cola, o unilaterales, y con g.l.= n – 1 = 10 – 1 = 9, y nivel de significación a = 0,01, encontrando un $t_c = 2,82$.

Como el valor de $t_e = 1,92 < t_c = 2,82$, no podemos rechazar H_0 a un nivel de significación de 0,01. Realizado el análisis mediante el programa informático SPSS nos aporta los siguientes resultados:

Tabla 14. Descriptivos

Estadísticos de muestras relacionadas					
	Grupos	**Media**	**N**	**Desviación típ.**	**Error típ. de la media**
	A = Control	5,50	10	2,013	0,63683
Par 1	B = Experimental	6,30	10	1,888	0,59722

Vemos que la media del grupo experimental es superior a la del grupo de control.

Tabla 15. Correlaciones

Muestras relacionadas				
		N	Correlación	Sig.
Par 1	Pun. Gr A y Pun, Gr B	10	0,774	0,009

La correlación entre las puntuaciones de ambos grupos es alta.

Tabla 16. Resultados de la prueba t

Prueba t para muestras relacionadas									
		Diferencias relacionadas					**t**	**gl**	**Sig. (bilateral)**
		Media	**Desvia. típica**	**Error típ. de la media**	**95% Intervalo de confianza para la diferencia**				
					Inferior	**Superior**			
Par 1	Pun G A Pun G B	-0,800	1,316	0,416	-1,74	0,141	-1,92	9	0,087

Los resultados obtenidos con el programa informático SPSS confirman los que ya habíamos obtenido aplicando la prueba directamente. La significatividad de alfa (α) es de 0,087, lo que equivale a su correspondiente nivel de confianza del 91,3%, no alcanzando el nivel de significación del 0,05, que es el aceptable científicamente para poder rechazar la hipótesis de nulidad H_0.

Las reflexiones sobre los resultados aconsejan que se amplíe la duración de la experiencia, y el tamaño de la muestra productora de datos para sacar conclusiones más seguras.

EJERCICIOS DE AUTOEVALUACIÓN

1. Defina los siguientes conceptos: Triangulación. Muestreo intencional. Credibilidad. Transferibilidad. Autenticidad.
2. Las calificaciones obtenidas por un grupo de 18 alumnos de 2.º de Bachillerato, en las pruebas de Matemáticas y Química (valoración de 10 a 1), han sido los siguientes:

   Calificaciones en Matemáticas: 4, 3, 6, 8, 7, 5, 4, 6, 5, 9, 7, 8, 4, 5, 4, 8,7,6.
   Calificaciones en Química: 4, 6, 8, 5, 4, 8, 7, 5, 7, 6, 8, 9, 7, 8, 9, 6,8,7.

Le pedimos que:

— Elabore la tabla de correlaciones.

— Calcule la media y la desviación típica.

— Calcule el coeficiente de variación.

— Calcule el coeficiente de correlación. Justifique la elección del coeficiente.

— Realice el diagrama de dispersión.

Utilice lo estudiado en el tema en la realización de evaluaciones de progamas educativos reales.

REFERENCIAS BIBLIOGRÁFICAS

Atlas.ti (2002). «Qualitative data Analysis». Scientific Software Development GmbH. Revised in 2017.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Boston: Houghton Mifflin.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis for field settings.* Chicago: Rand McNally.

García Jiménez, E., Gil Flores, J. & Rodríguez González, G. (1994). «Análisis de datos cualitativos en la investigación sobre la diferenciación educativa». *Revista de Investigación Educativa,* 23 (179-213).

Kerlinger, F. N. (1973) *Foundations of Behavioral Research.* 2nd edition. Holt, Rinehart and Winston.

Martínez Mediano, C. (Coord.). (2004). *Técnicas e Instrumentos de Recogida y Análisis de Datos.* Madrid. Unidades Didácticas de la UNED.5031UD01A01.

Miles, M. B. & Huberman, A. M. (1994). *Qualitative Data Analysis: an expanded sourcebook*. Newbury Park, CA, Sage,

Spss (2002) SPSS Inc. *Computer Sofware. Chicago: Illinois.* Programa informático para el análisis de datos estadísticos. Versión 11.

Taylor, S. J. & Bogdan, R. (1996). *Introducción a los métodos cualitativos de investigación*. Barcelona, Paidós.

ANEXO
Orientaciones para el estudio y la realización de un trabajo práctico

I. Prueba de evaluación continua

II. Orientaciones para la realización de un trabajo práctico sobre la evaluación de un programa educativo

1. Objetivos y resultados de aprendizaje.
2. Introducción.
3. Estructura del trabajo.
   - 3.1. Resumen.
   - 3.2. Introducción.
   - 3.3. El programa a evaluar.
     - 3.3.1. Fundamentos teóricos del programa.
     - 3.3.2. Objetivos del programa.
     - 3.3.3. Sujetos a los que se dirige el programa.
     - 3.3.4. El contexto.
     - 3.3.5. Recursos.
   - 3.4. La aplicación del programa.
     - 3.4.1. Estrategias y actividades.
     - 3.4.2. Utilización de los recursos.
     - 3.4.3. Resultados inmediatos a conseguir.
     - 3.4.4. Resultados a largo plazo.
   - 3.5. La evaluación del programa.
     - 3.5.1. Los objetivos de la evaluación, finalidad de la evaluación del programa.
     - 3.5.2. Funciones evaluativas.
     - 3.5.3. Modelo evaluativo elegido y Fases de la aplicación del modelo.

3.5.4. Metodología de la evaluación, técnicas de recogida y análisis de datos.
3.6. Resultados de la evaluación del programa.
3.6.1. Presentación y valoración de los resultados.
3.6.2. Conclusiones, discusión, valoración y propuestas de mejora.
3.7. Aplicación de los estándares de la evaluación de programas.
3.7.1. Aplicación de las normas a la evaluación.
3.7.2. Valoración del cumplimiento de los estándares.
3.8. Referencias bibliográficas y sitios webs.
3.9. Anexos.

4. Recomendaciones de estilo para la elaboración de trabajos de acuerdo con las normas APA.

5. Orientaciones para la realización del trabajo práctico y criterios para su evaluación.

Referencias bibliográficas.

I. PRUEBA DE EVALUACIÓN CONTINUA

Evaluación de Programas
Grado en Pedagogía

Alumno: nombre y apellidos:

Preguntas de la prueba

— Conteste con precisión, brevedad y suficiencia.
— Conteste directamente sobre el documento, en el mismo formato de letra y espacio.
— La longitud de la PEC, una vez contestada, debe estar entre 10 y 14 páginas. Las preguntas de la PEC varían de un curso a otro.

Tema 1. Concepto y funciones de la evaluación de programas

1. Conceptos de Evaluación normativa y Evaluación criterial.
2. Concepto de evaluabilidad y su finalidad.

Tema 2. M. Scriven. La evaluación como ciencia de la valoración

3. Importancia de la evitación de los prejuicios en la evaluación de programas.
4. Evaluación formativa y sumativa y su aplicación en la evaluación de programas.
5. Concepto de metaevaluación y su finalidad formativa y sumativa.

Tema 3. El modelo Cipp de D. Stufflebeam: La evaluacion orientada a la toma de decisiones para la mejora y el rendimiento de cuentas

6. Concepto de evaluación de programas de Stufflebeam. Valoración del mismo.

7. Haga una tabla de síntesis del modelo CIPP, sobre los cuatro tipos de evaluación, las decisiones a las que sirve y los instrumentos para la recogida de información que utiliza.

Tema 4. El modelo de evaluación de programas educativos de Ramón Pérez Juste

8. Concepto de evaluación de programas de Pérez Juste.
9. Definición de los criterios de adecuación o pertinencia, eficacia, eficiencia y satisfacción en relación con los programas educativos y su evaluación.

Tema 5. Procedimientos para la realización de evaluaciones de programas educativos. Serán utilizados en la parte orientada a la práctica de esta prueba.

10. Sobre un programa que conozca, elabore un diagrama en el que describa las necesidades o razones del programa, problemas que pretende resolver, sus metas y objetivos, los resultados que afirma pretender conseguir el programa, los factores personales, de recursos y de contexto necesarios y aconsejables para que el programa pueda conseguir sus objetivos, las estrategias y actividades que debe realizar el programa en su aplicación y los resultados a conseguir, orientados al logro de las metas finales del programa y su impacto.
11. Indique, para cada elemento del diagrama anterior, la metodología y técnicas de recogida de información a utilizar para su evaluación, comprobación de su funcionamiento y resultados. Muéstrelo mediante una tabla.

Tema 6. Las normas o estándares de calidad para la evaluación de programas educativos del Joint Committee

12. Haga una valoración de cada uno de los grupos de normas para juzgar la calidad de las evaluaciones. Mencione dos normas por cada grupo y justifique su elección

Supuesto práctico. Realización de la evaluación de un programa educativo

Póngase en la situación de que en su centro educativo se está aplicando un programa para la mejora del aprendizaje del inglés de los alumnos del primer curso de Educación Secundaria Obligatoria.

Como persona con la debida formación, la Dirección del centro le ha encargado la evaluación del nuevo programa, con especial incidencia en de la eficacia del programa.

Sobre tal base se le pide:

a) Precise los objetivos que se plantea para tal evaluación

b) Elija un modelo o propuesta evaluativa y justifíquela ante la dirección y el claustro de profesores.

c) Decida la metodología a utilizar, en particular, en relación con:

   i. Contenidos o aspectos fundamentales sobre los que versará la evaluación.

   ii. Fuentes productoras de datos.

   iii. Técnicas adecuadas para la recogida.

   iv. Organización y tratamiento de los datos.

   v. Utilización de la información.

d) Emita una valoración sobre las posibilidades y limitaciones de este tipo de evaluación.

Debe fijarse bien en el enunciado y responder de modo concreto y preciso a todas las preguntas, evitando dar respuestas en exceso genéricas.

II. ORIENTACIONES PARA LA REALIZACIÓN DE UN TRABAJO PRÁCTICO SOBRE LA EVALUACIÓN DE UN PROGRAMA EDUCATIVO

1. Objetivos y resultados de aprendizaje

Con la información recogida en este anexo pretendemos facilitar al estudiante enfrentarse a la realización de trabajos reales de evaluación de programas educativos. El esquema propuesto para la realización de la evaluación de un programa le ayudará a realizarlo paso a paso.

2. Introducción

El trabajo práctico consistirá en la «EVALUACIÓN DE UN PROGRAMA EDUCATIVO». Su finalidad es aplicar los contenidos que se estudian en la asignatura de Evaluación de Programas. Su realización ayudará al estudiante a comprender e integrar los contenidos teóricos desarrollados en el programa de la asignatura, a lo largo del curso, y a plasmar la comprensión de los mismos en un proyecto orientado a la práctica, concretado en la evaluación de un programa educativo real. Es interesante la realización de trabajos prácticos de aprendizaje con la finalidad de comprobar como la teoría tiene su eficacia en la realización de la evaluación de un programa real.

Conviene decir que por programa educativo entendemos todo programa que tiene finalidad de formar, educar, en cualquier ámbito, ya sea en contextos educativos formales, no formales, en el ámbito de la empresa, o en cualquier otro.

Entre los posibles programas a evaluar:

— Programa para la mejora de la lectura comprensiva en primero de la ESO.

— Programa para el desarrollo de estrategias colaborativas en el trabajo en grupo en segundo de la ESO o cualquier otro nivel y curso.

— Programa para el desarrollo del aprendizaje autónomo en cuarto de la ESO.

— Programa para la mejora de la expresión oral y escrita para sexto de Educación Primaria.

— Programa para la incorporación de las TICs para el aprendizaje en Educación Primaria.

— Programa de estrategias de aprendizaje para cuarto de Educación Primaria o cualquier otro curso.

— Programa de resolución de conflictos para la ESO.

— Programa para la prevención del acoso escolar en primero de ESO.

— Programa de desarrollo de la inteligencia emocional en Educación Primaria.

— Programa de desarrollo de competencias para el aprendizaje a lo largo de la vida en Educación Primaria, y Secundaria.

— Programa para la mejora del aprendizaje de las Matemáticas a través de las TICs para Primero de la ESO.

— Programa para la mejora del aprendizaje del inglés de los alumnos del primer curso de Educación Secundaria Obligatoria.

— Programa de desarrollo de competencias básicas en Educación Primaria.

— Programa de formación por competencias en la Formación Profesional.

— Programa de formación por competencias en la Educación Universitaria.

— Programa de desarrollo de competencias para el aprendizaje a lo largo de la vida en Educación Universitaria.

— Programa para el desarrollo de la formación en contenidos específicos en la empresa.

— Programa para el desarrollo de la formación en nutrición en los centros educativos.

— Programa para la formación en lenguaje de signos en los centros educativos.

Estas son algunas propuestas, pudiendo ampliarse de acuerdo con los intereses de cada persona y contexto. Como digo en otros apartados, parta la realización de la evaluación de un programa con fines de aprendizaje, no es conveniente comenzar con un programa institucional, como «El plan de formación del profesorado de la Comunidad Autónoma de...», «El uso de bibliotecas escolares en la Comunidad Autónoma de...» dado que estos programas requieren de equipos de profesionales para su evaluación. Para aprender hay que comenzar con un programa más breve y sencillo con el que estemos familiarizados.

3. Estructura del trabajo

3.1. Resumen

El resumen debe mostrar una síntesis completa de lo que se ha realizado en el trabajo. Los motivos, los objetivos, la metodología, los resultados. En torno a 300 palabras, de modo similar a como se realiza en los artículos de revistas científicas.

Palabras clave. Palabras claves, debajo del resumen. es una expresión, no solo una palabra, por ejemplo: Evaluación de programas, Investigación evaluativa, Nuevas tecnologías, Educación secundaria obligatoria. Entre 3 y 5, y deben ser representativas del trabajo. Resumen y palabras clave informan al lector de la esencia del trabajo.

También se puede poner el resumen y las palabras clave además de en español en inglés que serían Abstract y *Key words*.

Hacer esto en un trabajo práctico de aprendizaje ayuda a prepararse para luego realizar publicaciones en artículos científicos.

3.2. Introducción

Motivación del estudio, problema de estudio y justificación de su importancia, de modo breve, no más de medio folio. Es la expresión de motivos se explica por qué he decido evaluar este programa, argumentado en las necesidades sociales, de motivación personal, del entorno en el que trabajo, etc.

3.3. El programa a evaluar

Algunas preguntas de los alumnos sobre las «Características del programa educativo a evaluar» son del tipo siguiente: ¿es posible evaluar un programa de formación de profesorado que se está aplicando en mi comunidad? Y mi respuesta es que en una asignatura de aprendizaje no se puede comenzar con la evaluación de un programa institucional. Hay que comenzar con un programa sencillo, pequeño, con el que estemos familiarizados.

Otras preguntas que me hacen los alumnos son: ¿podría evaluar un programa deportivo? Mi respuesta es que si tiene fines de formación, de aprendizaje, se puede evaluar. Otros tipos de programas son acerca de una determinada aplicación informática. Mi respuesta es que sí, todo lo que suponga formación, aprendizaje, desarrollo, tiene fines educativos, luego es un programa de interés para evaluar dentro de esta asignatura de Evaluación de programas. Cualquier programa con fines de formación, de actualización, de nuevos aprendizajes, es un trabajo educativo.

3.3.1. Fundamentos teóricos del programa

Explicación breve del programa a evaluar. Podría ser un programa sobre la «Utilización de las tabletas electrónicas para el aprendizaje de las asignaturas de Ciencias Naturales y Ciencias Sociales en la Enseñanza Secundaria Obligatoria». También podría ser para sensibilizar a los estudiantes y a sus familias sobre la gravedad del acoso escolar. Si se trata de un programa para la sensibilización y reducción del acoso escolar, tendríamos que basarnos en los hechos que se dan de acoso escolar en los centros educativos y las graves consecuencias que pueden tener, y a partir de ello concretar un programa con objetivos estrategias, plan de actuación y de comprobación de sus logros y resultados de un modo claro, concreto, evaluable. Podría ser sobre la importancia de una alimentación sana, saludable, que reduzca posibles futuros problemas y enfermedades. Podría ser sobre el desarrollo de competencias para aprender a aprender y de ese modo incidir en las competencias para el aprendizaje a lo largo de la vida.

Sobre el programa a evaluar —la cuestión o problema de estudio—, se debe aportar fundamentación teórica y de programas similares, aportando referencias o trabajos sobre el tema, utilizando los apartados claves que demande el apartado. La información debe fluir de modo lógico.

3.3.2. Objetivos del programa

Los objetivos del programa deben estar formulados de modo claro, razonable, justificados por la fundamentación previa, por las necesidades que pretende atender. Si se trata de un programa para potenciar el uso de las tabletas electrónicas, deben concretarse para qué se introduce este elemento en las asignaturas de Ciencias Naturales y de Ciencias Sociales, con la idea de que su uso facilita el acceso a información complementaria a la que se recogen en los textos de la asignaturas, para ampliar la información, para búsqueda de recursos de aprendizaje como videos especificos sobre los temas del programa de estudio, pero también para la elaboración de trabajos en formato electrónico. Con ello se pretende reforzar los objetivos específicos del propio programa de estudio, así como el desarrollo de competencias clave como el uso de las tecnologías de la información y de la comunicación, el trabajo en equipo, la presentación de trabajos, lo que redundara en el beneficio de la preparación de los alumnos.

3.3.3. Sujetos a los que se dirige el programa

En este apartado hay que explicar a qué grupo de alumnos se dirige el programa, indicando alguna característica de los mismos.

3.3.4. El contexto

Explicar el contexto en el que se desarrolla el programa, el curso, las características del centro educativo. No es necesario una información muy detallada. Tampoco es necesario mencionar la zona en la que se encuentra el centro, y si se hace de modo breve, ya que es un trabajo de prácticas en donde el programa, y especialmente su evaluación e informe final, es lo importante, en este caso.

3.3.5. Recursos

Además de la tableta electrónica, qué otros recursos serían necesarios utilizar. Por supuesto el propio programa, el personal que lo aplica, la organización del centro al servicio del buen desarrollo de los programas.

3.4. La aplicación del programa

Para explicar este apartado debemos conocer las características de los programas educativos y también el programa concreto que vamos a evaluar.

3.4.1. Estrategias y actividades

Mencionar, explicar y justificar las estrategias necesarias para que el programa sea aplicado con éxito. Por ejemplo, sería conveniente realizar una sesión para que el profesor, o el experto en informática mostraran las posibilidades que ofrece para el estudiante la tableta electrónica y enseñar dichas posibilidades de uso a los alumnos. Trabajo individual y grupal de los estudiantes utilizando recursos tecnológicos para la elaboración, por ejemplo de una exposición mediante un vídeo elaborado por los alumnos.

3.4.2. Utilización de los recursos

Por ejemplo, referidos al tiempo de utilización de la tableta en el aula y fuera del aula.

3.4.3. Resultados inmediatos a conseguir

Lo que se pretende conseguir, mejora en el uso de las tabletas con objetivos de aprendizaje, mejora en el aprendizaje de los contenidos sobre el programa de estudio de Ciencias Naturales. Los resultados deben ser coherentes con los objetivos pretendidos por el programa.

3.4.4. Resultados a largo plazo

Si se trata de un programa que utiliza las TICs, a largo plazo se pretende desarrollar la capacidad de búsqueda de información a través de internet, así como el desarrollo de habilidades para la realización de trabajos mediante el uso de los recursos tecnológicos.

Si se trata de un programa de mejora nutricional, el impacto sería que los alumnos aprendieran la importancia de alimentarse de un modo saludable. Y si en el programa se ha conseguido implicar a los padres, el impacto será mayor, dado que se habrá podido conciencias a los padres de la importancia de realizar comidas saludables, y su aplicación en la vida real tendrá un impacto superior al que pueda conseguirse solo mediante la concienciación de los alumnos.

3.5. La evaluación del programa

Todo programa debe pretender conseguir unos objetivos, debe desarrollar y explicar cómo va a actuar para lograr dichos objetivos, de acuerdo con unos modelos o concepciones teóricas en las que se fundamenta. Por ejemplo, si se trata de un programa de enseñanza, deben estar explícitas, o implícitas, las concepciones de las teorías de aprendizaje que en él subyacen y que están al servicio del logro de los objetivos, formando parte de la estructura y coherencia interna del programa. Asimismo, debe tener una población a la que el programa se dirige, estando claramente identificada, y con unas expectativas en consonancia con lo que el programa ofrece. El programa, en este aspecto, debe definir indicadores de información hacia la población a la que se dirige, para informar de lo que ofrece, e identificar el conocimiento e interés de los grupos a los que atiende el programa.

Para su adecuada aplicación, el programa debe contener una estructura organizativa coherente para su desarrollo, y de manera explícita debe tener definidas unas funciones y describir unas estrategias y procedimientos para el logro de los objetivos que se propone conseguir. El programa, para su desarrollo, debe contar con unos recursos personales y materiales, que hagan posible la aplicación y desarrollo del programa.

El programa debe definir unos indicadores de ejecución, de modo que informe sobre cómo debe ser implementado el programa, y de ese modo facilitar comprobar si se aplica como se había planeado, y si se logran los resultados pretendidos, definiendo los resultados a conseguir. Debe concretar las acciones a realizar mediante los procesos, con una duración definida en el tiempo, señalando su comienzo y su final.

La evaluación está al servicio de objetivos educativos de calidad, para la formación integral de los estudiantes. La evaluación debe informar de la calidad de esos objetivos, de los programas a su servicio, con su estructura, estrategias, actividades, logros inmediatos que consigue y resultados generales y su impacto. Los objetivos de los programas y los elementos para su desarrollo son los criterios que orientan la evaluación. Y a partir de ellos se deben elaborar los indicadores para su evaluación. Por ejemplo, de acuerdo con las etapas del modelo de Pérez Juste para la evaluación de los programas educativos, los criterios de cada una de las etapas deben originar los indicadores para su evaluación. Las dimensiones en las que se centran en cada momento o etapa de la evaluación, y el objeto de la evaluación, señalan el objetivo u objetivos que debe conseguir, mediante un programa, de calidad —de acuerdo con los objetos/elementos en los que debe centrarse—, que responda a las necesidades para las que se crea, y que cuente con los suficientes medios, recursos, personas, para que su aplicación y la consecución de sus metas sea viable. Entonces, evaluar la etapa inicial requiere elaborar indicadores representativos, significativos de lo que esta etapa presenta.

Y así en todas los momentos o etapas del proceso evaluación en el modelo de evaluación de programas de Pérez Juste. Y todo ello adaptado a las características del programa que vamos a evaluar —y su contexto—.

3.5.1. Los objetivos de la evaluación, finalidad de la evaluación del programa

Qué finalidad evaluativa se tiene: conocer las características del programa y evaluarlas, valorar si el uso de las tabletas electrónicas consigue los resultados pretendidos. Aportar información sobre el programa y los resultados que consigue.

3.5.2. Funciones evaluativas

Evaluación formativa, a lo largo de la aplicación del programa y sumativa, para comprobar los resultados que aporta el programa, en relación con los objetivos que pretendemos y si consigue otros resultados no pretendidos, positivos y negativos. Funciones administrativas, para informar al claustro de profesores y a los padres de los resultados que consigue. El potencial que aporta el uso de las tecnologías.

3.5.3. Modelo evaluativo elegido y Fases de la aplicación del modelo

Justificar por que se opta por un determinado modelo evaluativo, o por elementos de varios modelos. Fundamentarlo. Planificación de la evaluación de acuerdo con el modelo en sus distintas fases. Por ejemplo, el modelo CIPP debe reflejar el análisis de necesidades en su contexto, el análisis de la planificación para su aplicación, el análisis de su aplicación y resultados que va consiguiendo e información acerca de si la aplicación del programa ha exigido modificar el plan previsto y por qué. Evaluación de resultados que consigue.

3.5.4. Metodología de la evaluación, técnicas de recogida y análisis de datos

Siguiendo, por ejemplo, el modelo CIPP, en el análisis de contexto, el análisis de necesidades puede basarse en peticiones de los padres, en manifestaciones de los profesores, en propuestas de ambos grupos para potenciar el uso de las tabletas electrónicas con fines de aprendizaje. Ello lleva a el desarrollo de un programa de aplicación de las tablas, a elaborar un plan de evaluación, y a aportar información para la evaluación de entrada, a través del análisis por expertos de compañeros sobre el programa y el plan de aplicación, y las estrategias planificadas, las actividades, y el sistema de evaluación. Ello permitirá contar con el diseño y recursos para la aplicación del programas, que nos permitirá elaborar instrumentos para recoger información sobre el proceso de aplicación y resultados que consigue, y finalmente recoger

información acerca de los resultados que ha conseguido el programa en relación con lo que pretendía.

En relación con la metodología a utilizar, hay que precisar los:

— Los contenidos o aspectos fundamentales sobre los que versará la evaluación.

— Las fuentes productoras de datos.

— Las técnicas adecuadas para la recogida.

— La organización y tratamiento de los datos.

— La utilización de los resultados de la información obtenida en la evaluación.

Los contenidos de la evaluación sería los específicos del programa, por ejemplo si se trata de programa de mejora del inglés, tendríamos que fijarnos en su estructura, su aplicación y su evaluación; y si la evaluación del programa va a priorizar la evaluación de la eficacia del programa en el logro de sus objetivos, se debe centrar especialmente en valorar la eficacia de dicho programa, aunque sin perder de vista que la evaluación de programas siempre es holística, sistémica, las partes se influyen mutuamente, y entre todas se consiguen un cuadro global del programa, en sí mismo, en su aplicación y en los resultados que consigue. Y estos deben ser valorados críticamente y en relación con el contexto en el que se desarrollan.

Si, por ejemplo, estamos centrados en la evaluación de proceso, sobre un programa de mejora del aprendizaje de las matemáticas, recogeríamos información del aprendizaje de los estudiantes mediante la realización de una prueba de aprendizaje, la técnica sería una prueba de examen pudiendo ser una prueba objetiva, o de ensayo, sobre conceptos y resolución de problemas.

Otra fuente productora de datos serían los propios profesores que están aplicando el programa, y que tras su aplicación, por ejemplo, al finalizar el primer trimestre, o el primer mes, deciden tener una reunión con la dirección y mediante un grupo de discusión debaten acerca de los resultados que se están consiguiendo, junto con los problemas que han ob-

servado por no disponer de algún recurso que argumentan que es importante para que el programa consiga los objetivos pretendidos.

En la evaluación de resultados, una fuente productora de datos serían las calificaciones obtenidos por los alumnos, que aportarían información sobre la eficacia del programa en términos de rendimiento, de logro de los objetivos.

Otra fuente productora de datos, para valorar el éxito de un programa pueden ser los padres, por ejemplo, en relación con la aplicación de un programa para la mejora de la convivencia, para la resolución de conflictos, para la sensibilización y reducción del acoso escolar, por lo que se les debe preguntar a los padres, antes de aplicar el programa para recoger aportaciones importantes basadas en su experiencia con sus hijos en relación con estos temas, durante la aplicación para conocer algunos indicadores sobre lo que el programa está consiguiendo, y al final de determinados períodos, del primer trimestre, si observan cambios en los comportamiento de sus hijos, y al finalizar el curso, a través de una entrevista en grupo, entrevistas individuales con los tutores o mediante un cuestionario.

3.6. Resultados de la evaluación del programa

3.6.1. Presentación y valoración de los resultados

Cómo se organizan y se presenta la información recogida, los datos, pues mediante tablas de datos, síntesis de información textual, y se realizarían análisis estadísticos de los datos, eso es el tratamiento de los datos, los análisis estadísticos, que se presentaría mediante gráficos, tablas de resúmenes de descriptivos, comparación de grupos, pruebas de hipótesis cuando fuera necesarios... y se mostraría todo ello en un informe adaptado a los diferentes grupos de interés.

La presentación de los resultados es el apartado más técnico del trabajo. Los resultados deben presentarse utilizando los subapartados que sean necesarios. Describir los análisis realizados. Utilizar tablas para resumir datos (por ejemplo, con medias y desviaciones típicas, valores de las pruebas, nivel de significación). No repetir la información de las ta-

blas en el texto, evitar ser redundante. Utilizar figuras para visualizar información completa. Situarlas al final de la página. Mencionar la tabla y su número en el texto. Los títulos de las tablas y de las figuras deben ser significativas de lo que representan, y seguir numera separada, tablas y figuras. El título de la tabla sobre ella, el de la figura bajo ella, numerados. Debemos mostrar de una manera sistemática los resultados aportados por la evaluación del programa.

Resultados de rendimiento y de satisfacción de los alumnos, de los profesores y padres sobre el uso de las tabletas electrónicas y de sus resultados. Información sobre los resultados, interpretación y valoración de los mismos.

3.6.2. Conclusiones, discusión, valoración y propuestas de mejora

Las conclusiones forman parte de los resultados y podrían considerarse un resumen de los resultados clave introductorios de la discusión y valoración crítica de los mismos. A partir de los resultados y su interpretación valoramos los pros y los contras y aportamos valoraciones, recomendaciones, sugerencias.

El propósito del apartado de discusión es resumir, interpretar, evaluar, valorar críticamente los resultados y sus implicaciones. No se pueden reiterar de modo literal los resultados. Se trata de una reflexión crítica, junto con una interpretación e implicación de los resultados del estudio.

Debe comenzar con una breve introducción del objetivo del estudio, valorando si las estrategias utilizadas mejoran los resultados de estudios previos, y articular, mencionar de modo organizado sus fortalezas, sus aportaciones. Se puede proporcionar un breve resumen del estado de la cuestión previa para a continuación informar si nuestra propuesta, el programa aplicado, contribuye a la mejora en línea con nuestros objetivos al aplicarlo, con los objetivos del programa, e interpretar y valorar si ha conseguido mejoras más allá de las esperadas y por qué. Hacer consideraciones sobre el significado práctico de los resultados para su aplicación en futuros curso o recomendar su utilización en situaciones similares. Incluir consideraciones sobre las limitaciones del estudio con

respecto al propio programa y a su aplicación a los recursos a los instrumentos de recogida de información para la evaluación en sus diferentes fases, y acerca de la validez interna, la coherencia y adecuación de las metodologías e instrumentos de recogida de información utilizados. Reconocer las limitaciones, posibles errores metodológicos, debilidades del trabajo, cuidando no destacar solo lo positivo del trabajo, ni tampoco solo lo negativo.

Emitir una valoración sobre las posibilidades y limitaciones de este tipo de evaluación. Señalar futuras direcciones de nuevos trabajo. Discutir su aplicación a contextos reales, y su relevancia para la práctica y el cambio necesario. Ser preciso, exacto.

Recomendaciones de mejora. Basándonos en toda la información anterior hacemos recomendaciones para mejorar el programa y su aplicación. Reducir el programa o ampliarlo. Y también para expandir su aplicación a otras materias, información a los padres, el claustro y compañeros.

3.7. Aplicación de los estándares de la evaluación de programas

3.7.1. Aplicación de las normas a la evaluación

Se deben utilizar las indicaciones contenidas en el tema de los estándares y aplicar la tabla que aparece en el tema, con las sugerencias que se recogen en el mismo.

3.7.2. Valoración del cumplimiento de los estándares

Los estándares, integrados por cinco grupos de estándares o categorías, que contienen 30 normas, orientan acerca de cómo se ha evaluado el programa, y además utilizándolos desde el principio ayudan al diseño del programa y su evaluación. Utiliza una tabla para valorar cada una de las normas mediante una valoración de 1 a 5, o de 1 a 4, con una primera opción de si se ha aplicado o no, y si se ha hecho en qué grado se ha podido explicar, comentando las valoraciones más llamativas, y si se han podido aplicar todas las normas o alguna de ellas no era posible o adecuado su aplicación, explicando los motivos.

3.8. Referencias bibliográficas y sitios webs

Las citas de la bibliografía y páginas web consultadas deben mencionarse en el texto del trabajo siguiendo las normas de la American Psycgological Association (APA) que son las que más se utilizan en Educación y Ciencias Sociales. Por lo tanto se debe aplicar a lo largo del trabajo. Y luego todas las referencias bibliográficas y de páginas web utilizadas que están reflejadas en el texto deben incluirse en las «Referencias» al final del trabajo.

Nunca se debe omitir en las referencias finales una cita reflejada en el texto del trabajo. Y tampoco se deben incluir referencias no citadas en el texto. Las referencias deben ir en orden alfabético del apellido del primer autor, con la primera letra en mayúscula. Después del autor o autores se pone entre paréntesis el año de publicación. Luego el título en itálica, si es un libro, después el nombre de la ciudad donde se ha publicado y seguido del nombre de la editorial; cuando se trata de un artículo, el título del artículo va en letra normal y nombre de la revista en itálica. Después se pone el número de la revista en itálica, y páginas en normal. Autores, fecha, título, ciudad separados por punto. Ciudad y editorial separados por dos puntos. Título de revista, número y páginas, separadas por comas. Para mayor información consultar APA style Guide, en http://www.apastyle.org. También se recomienda leer artículos científicos con el fin de profundizar en el tema del programa que se evalúa y así mismo el modo como se cita en el texto y en las referencias al final del trabajo o artículo.

3.9. Anexos

Incluir anexos en un trabajo es «optativo», en el sentido de que no tiene por que estar en todos los trabajos. Dado que los trabajos de aprendizaje suelen tener limitación de espacio, y así mismo los artículos de revistas, se aconseja incluir lo que iría en los anexos en el propio texto del trabajo, en su apartado correspondiente, teniendo en cuenta el espacio del que se dispone. Por ejemplo, un cuestionario iría en el apartado de metodología, en el subapartado de instrumentos.

4. Recomendaciones de estilo para la elaboración de trabajos de acuerdo con las normas APA

Los títulos y subtítulos de los apartados y subapartados, cada uno en su mismo formato.

Sujeto y verbo siempre congruente en número y género. Los verbos del texto siempre en pasado excepto en los resultados, conclusiones y discusión que debe ir en presente. Se debe utilizar la voz activa, siendo siempre preferible a la voz pasiva.

La citas directas se deben poner cuando sea crucial, entre doble «marca de cita» si es inferior a 40 palabras y cuando las supera debe ponerse en un párrafo separado, sin marcas de cita.

Las abreviaturas o acrónimos, la primera vez es completo y la abreviatura entre paréntesis. Sucesivas veces poner solo la abreviatura.

En las referencias, si los autores van en el texto, ponerlos todos hasta 6, y luego poner *et al.,* el penúltimo autor va unido al último con la palabra y; cuando los autores van entre paréntesis, en lugar de 'y' van unidos con el símbolo &. El año entre paréntesis, indicando el número de página, por ejemplo: (2011, p. 89, o pp. 89-92).

Cuando se ponen series, si son letras, entre paréntesis completo (a), si son números sólo el número 1. seguido de punto. (Cash, 2009).

5. Orientaciones para la realización del trabajo práctico y criterios para su evaluación

El trabajo práctico consistirá en la «EVALUACIÓN DE UN PROGRAMA EDUCATIVO». Su finalidad es aplicar los contenidos que se estudian en la asignatura de Evaluación de Programas**.** Su realización ayudará al estudiante a comprender e integrar los contenidos teóricos desarrollados en el programa de la asignatura, a lo largo del curso, y a plasmar la comprensión de los mismos en un proyecto orientado a la práctica, concretado en la evaluación de un programa educativo real.

La valoración y calificación del trabajo práctico se realizarán de acuerdo con las orientaciones de calificación que se indican a continuación.

El trabajo deberá contener los siguientes apartados:

	Puntuación
Resumen. Del trabajo, reflejando sus distintos apartados, máximo de 300 palabras. **Palabras claves.** Un máximo 5 expresiones o palabras claves representativas del trabajo realizado. 1. **Introducción.** Presentación razonada de los motivos personales de la elección del programa a evaluar.	0,5 si el resumen refleja lo que se pide. 0,25 palabras claves del trabajo. 0,25 a la introducción refleja lo que se pide. Total 1
2. **El programa.** Basándose en el análisis del programa a evaluar, que debe reunir como principal característica que tenga fines educativos. Nombre del programa sobre el que realiza el trabajo. 2.1. Objetivos del programa. 2.2. Fundamentos teóricos del programa. 2.3. Sujetos a los que se dirige el programa. 2.4. El problema que pretende subsanar, prevenir o mejorar. 2.5. Contexto en el que se aplica. 2.6. Duración de la aplicación del programa.	0,5 Formulación correcta de los objetivos. 0,5 Fundamentos teóricos del programa. 0,5 Sujetos y el problema a solucionar y el contexto y duración Total 1,5
3. **La aplicación del programa.** Fundamentada en el análisis del propio programa. Información sobre cómo debe aplicarse el programa. 3.1. Recursos, incluida la formación que deben tener las personas que lo aplican. 3.2. Estrategias y actividades que debe realizar el programa para conseguir los resultados pretendidos. 3.3. Los resultados inmediatos que se esperan conseguir. 3.4. Resultados que se espera que consiga el programa a medio y largo plazo.	0,25 Información sobre recursos 0,75 Explicación estrategias y actividades 0,25 Resultados inmediatos 0,25 Su impacto Total 1,5

	Puntuación
4. **La evaluación del programa.** (Basada en los contenidos que se estudian en la asignatura de Evaluación de Programas. Se estudia en los temas 1, 2, 3, 4, 5, 6). Información argumentada sobre cómo se va a evaluar el programa, su concepción teórica —de evaluación de programas. 4.1. Los objetivos evaluativos que persigue con la evaluación del programa (Cualquier concepto de evaluación de programas lo recoge, y se estudia en los temas 1, 2, 3, 4, 5) 4.2. Las funciones evaluativas que pretende, (temas 1, 2, 3, 4, 5) 4.3. El modelo evaluativo al que se circunscribe, o síntesis de varios, y explicación y justificación de su elección. (temas 2, 3, 4, 5) 4.4. Fases de la evaluación del programa (se estudia en los temas 2, tipos de evaluación CIPP, 3, momentos o fase, Pérez Juste, 5, Procedimientos, modelo lógico), 4.5. Metodología de evaluación a utilizar, incluidas las técnicas de recogida de datos y las técnicas de análisis de datos (Se estudia en los temas 2, 3, 4, 5, 7 y 8).	0,5 Objetivos evaluativos del programa 0,5 Funciones de la evaluación del programa 1 El modelo 0,5 aplicación del modelo en sus fases o momentos, o tipo de evaluación. 0,5 La metodología a aplicar en cada fase. Total 3
5. Resultados de la evaluación del programa. (Se estudia en los temas 2, 3, 4, 5, 7 y 8 y anexos) 5.1. Presentación de los resultados de la evaluación del programa, en coherencia con los motivos y objetivos que llevan a aplicar el programa. 5.2. Valoración de los resultados que ha conseguido el programa. 5.3. Conclusiones y recomendaciones de mejora. Si se puede, recoger información del funcionamiento del programa mejor, debiendo incluir los resultados de la recogida de datos.	0,5 Resultados de la evaluación en coherencia con los objetivos de la evaluación. 0,25 Valoración de los resultados y conclusiones 0,25 Recomendaciones de mejora. Total 1

	Puntuación
6. **Aplicación de las Normas del Joint Committee a la evaluación del programa realizado.** (Se estudia en el tema 6). 6.1. Aplicación de las normas a la evaluación realizada mediante una tabla. 6.2. Valoración sobre los estándares cumplidos en la evaluación de programa realizado.	0,5 Aplicación correcta de las Normas. 0,5 Comentarios, valoración de los resultados de su aplicación Total 1
7. **Referencias bibliográficas y sitios webs.** Siguiendo las normas de la American Psycological Association APA, en **http://www.apastyle.org** En la página Plan de trabajo se facilita un resumen en español de las mismas. 7.1. Siguiendo las normas para la citación de la información que se toma de libros, artículos, páginas webs, a lo largo del trabajo. 7.2. Las referencias al final del trabajo.	0,5 Aplicación de las Normas de la APA a lo largo del trabajo. 0,5 presentación correcta de las referencias bibliográficas. Total 1
La presencia de cada uno de los apartados no da lugar a su máxima puntuación. La máxima puntuación se da cuando el apartado está bien desarrollado, argumentado, razonado, bien redactado y bien presentado. La adecuada fundamentación en los contenidos del programa de la asignatura es un requisito necesario para la realización del trabajo. A medida que esta corrección y adecuación decrece la puntuación lo hace también. La longitud del trabajo tendrá un mínimo de 15 páginas y un máximo de 20, incluidos los datos del estudiante que realiza el trabajo, y su desarrollo, incluidos posibles anexos que se recomiendan evitar en este tipo de trabajos, debiendo incluirse todo lo importante en el propio texto del trabajo; escrito en tipo de letra Times, 12 puntos, espacio entre líneas 1,15. La copia y el plagio serán penalizados.	

REFERENCIAS BIBLIOGRÁFICAS

American Psycological Association APA, en http://www.apastyle.org.

CASH, T. F.. (2009). *Caveats in the proficient preparation of an APA-Style resarch manuscript for publiction.* Body Image, 6, 1-6.

En las últimas tres décadas, los modelos de evaluación de programas con objetivos de mejora y de aseguramiento de la calidad para atender a las necesidades y problemas educativos y sociales, se han convertido en elementos esenciales para dirigir los esfuerzos por la mejora continua de los programas y de las organizaciones. Todo parece ser motivo de evaluación: el aprendizaje de los alumnos, el desempeño docente, el diseño y desarrollo del currículo y de los programas educativos, la organización de centros, entre otros. Los propósitos son muy diversos. Se puede evaluar para conocer, para valorar, para promocionar, para mejorar, pero también para rendir cuentas ante la sociedad, para acreditar o para decidir el futuro de un programa educativo. El salto cualitativo de la evaluación es su orientación hacia la mejora, en su función formativa, de modo que, entre todos los propósitos que pueden inspirar la tarea evaluativa, merece la pena destacar el de la mejora, ya sea referida a la formación del estudiante, del docente, de la organización, o de los sistemas educativos.

Los procedimientos que utiliza la evaluación están al servicio de los objetivos de las intervenciones y a ellos deben adaptarse. La realidad social es compleja, también sus necesidades y los programas que se desarrollan para atenderlas. También son complejos los contextos en los que se aplican, los resultados que producen y su impacto, incluidos los efectos por transferencia a otras situaciones y personas. Esta complejidad hace necesaria la adecuación de los procedimientos para la evaluación, la utilización de las metodologías de investigación de un modo riguroso, adaptadas a los múltiples objetivos que abarca la evaluación en los contextos en los que los programas y sus aplicaciones se desarrollan. La evaluación siempre aportará información para la mejora, para introducir cambios y actualizaciones para adaptarse a las nuevas demandas de los avances que continuamente se producen en la sociedad.

UNED Editorial

Juan del Rosal, 14
28040 MADRID
Tel. Dirección Editorial: 913 987 521